

LAPORAN
PENELITIAN HIBAH KOMPETENSI



JUDUL PENELITIAN
MODEL ASSESSMENT TERSTANDAR BERBASIS *COMPUTER*
***MANAGEMENT INSTRUCTIONAL* UNTUK MENJAMIN KESETARAAN**
KUALITAS PENILAIAN SEBAGAI BASIS DATA PENENTUAN KELULUSAN
DALAM SISTEM UJIAN AKHIR NASIONAL DAN SNMPTN JALUR
UNDANGAN YANG BERKEADILAN

TIM PENELITIAN:

Dr. Dadan Rosana, M.Si. (NIDN. 0002026904)
Dr. Sukardiyono, M.Si. (NIDN.0016026603)

UNIVERSITAS NEGERI YOGYAKARTA
Oktober 2015

HALAMAN PENGESAHAN

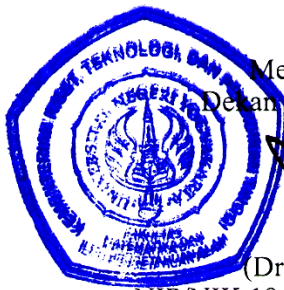
Judul : MODEL ASSESSMENT TERSTANDAR BERBASIS
CMI (COMPUTER MANAGEMENT
INSTRUCTIONAL) UNTUK MENJAMIN
KESETARAAN KUALITAS PENILAIAN DI
SEKOLAH SEBAGAI BASIS DATA PENENTUAN
KELULUSAN SISTEM UJIAN AKHIR NASIONAL
YANG BERKEADILAN

Peneliti/Pelaksana

Nama Lengkap : Dr. DADAN ROSANA M.Si.
Perguruan Tinggi : Universitas Negeri Yogyakarta
NIDN : 0002026904
Jabatan Fungsional : Lektor
Program Studi : Pendidikan Ilmu Pengetahuan Alam
Nomor HP : 081392859303
Alamat surel (e-mail) : danrosana@uny.ac.id.

Anggota (1)

Nama Lengkap : Dr. SUKARDIYONO M.Si.
NIDN : 0016026603
Perguruan Tinggi : Universitas Negeri Yogyakarta
Institusi Mitra (jika ada)
Nama Institusi Mitra : -
Alamat : -
Penanggung Jawab : -
Tahun Pelaksanaan : Tahun ke 2 dari rencana 2 tahun
Biaya Tahun Berjalan : Rp 122.500.000,00
Biaya Keseluruhan : Rp 300.000.000,00



Mengetahui,
Dean FMIPA UNY

(Dr. Hartono)
NIP/NIK 196203291987021002

Yogyakarta, 7 - 11 - 2015

Ketua,

(Dr. DADAN ROSANA M.Si.)
NIP/NIK 196902021993031002



Menyetujui,
Ketua LPPM UNY

(Prof. Dr. Anik Ghufro)
NIP/NIK 196211111988031001

RINGKASAN

Permasalahan sekitar ujian akhir nasional (UAN/UN) sampai saat ini masih menjadi permasalahan utama dalam bidang pendidikan yang melahirkan banyak kontroversi, salah satunya adalah tentang cara penentuan kelulusan. Berkaitan dengan penentuan kelulusan telah diterbitkan Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2013 Tentang Kriteria Kelulusan Peserta Didik Dari Satuan Pendidikan dan Penyelenggaraan Ujian Sekolah/Madrasah/Pendidikan Kesetaraan dan Ujian Nasional. Pasal 6 ayat 3 mengungkapkan bahwa, Nilai akhir untuk penentuan kelulusan diperoleh dari gabungan nilai sekolah dari mata pelajaran yang diujikan secara nasional dan Nilai UN, yaitu dengan pembobotan 40% Nilai sekolah dari mata pelajaran yang diujikan secara nasional dan 60% dari Nilai UN. Permasalahan yang kemudian muncul berkaitan dengan hal ini adalah belum adanya kesetaraan kualitas assessment yang digunakan untuk penilaian di sekolah (semester 3 (tiga) sampai dengan 5 (lima) untuk tingkat SLTP/SLTA/SMK), sehingga belum dapat menjamin rasa keadilan karena perbedaan kualitas tes yang diberikan. Hal ini sangat mendesak untuk dicarikan solusinya, karena nilai sekolah digunakan juga dalam sistem penerimaan mahasiswa baru (SNMPTN) jalur undangan.

Penelitian menggunakan metode deskriptif kuantitatif. Hasil deseminasi penyetaraan item tes anchor dari Soal UAS yang digunakan bersama oleh 16 sekolah pada tahun ajaran 2014/2015, di sekolah mitra di Indonesia (6 SMANTB, 2 SMA Lampung, 2 SMAPalangkaraya, 2 SMADIY dan 4 SMA Jawa Barat) yang diambil secara purposive sampling (kualitas tinggi, sedang, rendah). Realisasi dari tujuan penelitian yang telah dilaksanakan pada tahun kedua ini adalah; (1) telah berhasil dikembangkan dan dideseminasikan model assessment terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian di 16 sekolah mitra se Indonesia sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan, (2) telah diterapkannya tes terstandar sesuai dengan kompetensi inti dari kurikulum 2013 yang dapat digunakan sebagai perangkat equiting (penyetaraan) dalam penyusunan tes di sekolah mitra di Indonesia sehingga dapat digunakan sebagai penentu kelulusan dalam sistem ujian akhir nasional yang berkeadilan, (3) telah dihasilkan Bank soal terstandar untuk digunakan dalam penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan.

Luaran penelitian yang telah dihasilkan sampai tahap ini adalah; (1) Software CMI Sistem Informasi Penilaian SMA (SIPSMA), (2) Manual user book untuk program CMI SIPSMA, (3) *Proceeding of International Conference On Research, Implementation And Education Of Mathematics And Sciences (ICRIEMS) 2015, Yogyakarta State University, 17-19 May 2015* (ISBN : 978-979-96880-8-8, lampiran 1), (4) *Proceeding of International Conference On Educational Research and Innovation (ICERI), Institute of Research and Community Services Yogyakarta State University, 6-7 May 2015* (ISSN: 2443-1753, dan publikasi jurnal nasional terakreditasi/ jurnal internasional; (1) *Aceptance* Jurnal Kependidikan (LPPM UNY), Jurnal Nasional Terakreditasi (2) proses review Jurnal Ilmu Pendidikan (Universitas Negeri Malang), Jurnal Nasional terakreditasi dan (3) Submit jurnal internasional di International Journal of Education and Learning (Canadian Centre of Science and Education).

Hasil analisis menunjukkan bahwa kualitas soal di masing masing wilayah memiliki tingkat kesukaran yang berbeda. Paket soal di wilayah Jawa Barat paling tinggi dibandingkan paket soal di wilayah lain. Persamaan linear equating yang terbaik terhadap skor akhir yang tidak merugikan peserta ujian adalah penyetaraan dari paket sukar ke paket mudah dengan persamaan; (1) $Y' =$

1,543 X – 3.976, untuk equating dari Paket A ke B (2) $Y' = 1,876 X - 14,342$, untuk equating dari Paket B ke C (3) $Y' = 1,443 X - 3,203$, untuk equating dari Paket C ke A , (4) $Y' = Y' = 1,384 X - 0.932$ untuk equating dari Paket D ke A , (5) $Y' = 1,214 X - 0,607$ untuk equating dari Paket A ke C, (6) $Y' = 1,229 X - 4,987$ untuk equating dari Paket B ke C, dan (7) $Y' = 1,256 X - 1,660$. untuk equating dari Paket B ke D.

. Penelitian ini juga mendeteksi ketidakwajaran sekor dengan korelasi person biserial (*person-fit statistic*) dan indeks kehati-hatian (*caution index*) dari Sato. Paket program yang digunakan untuk melakukan analisis butir adalah Quest, dengan elemen sentral *Rasch Model* (RM). Dari hasil analisis butir soal diperoleh bahwa *coution index* dari Sato terdeteksi 11,56 persen sample mempunyai ketidakwajaran, dan 18,72 persen sample yang diambil mempunyai tingkat ketidakwajaran yang tinggi.

PRAKATA

Alhamdulillah penelitian tahun ke dua tahap pertama dan laporan penelitian mengenai Model Assessment Terstandar Berbasis CMI (*Computer Management Instructional*) untuk Menjamin Kesetaraan Kualitas Penilaian Sebagai Basis Data Penentuan Kelulusan dalam Sistem Ujian Akhir Nasional dan SNMPTN Jalur Undangan yang Berkeadilan ini dapat terlaksana dan tersusun. Hasil penelitian ini diharapkan dapat memenuhi standard penelitian baik dari segi metodologi maupun mekanismenya, sehingga ke depan penelitian ini dapat didanai dan dikembangkan lebih lanjut.

Sebagai sebuah penelitian tahun kedua, sedang berjalan tentu saja masih terdapat langkah-langkah yang harus diselesaikan untuk mendapatkan data, perhitungan, dan analisa agar dapat dicapai tujuan yang diharapkan. Data dan analisa yang tersaji merupakan data yang bersifat sampel nasional dan dapat dijadikan referensi awal pada 16 sekolah mitra tempat penelitian dilakukan, mudah-mudahan sudah dapat mewakili karakteristik ke-Indonesiaan.

Model Assessment Terstandar Berbasis CMI (*Computer Management Instructional*) yang telah dikembangkan untuk tingkat SMA, sehingga pengembangannya masih memerlukan deseminasi lebih luas, mengingat masih sangat beragamnya kualitas sekolah di wilayah negara kesatuan Republik Indonesia. Kami berharap sekecil apapun kontribusi yang dapat diberikan tetap dapat memberikan manfaat bagi penelitian-penelitian berikutnya. Tidak lupa terima kasih kepada DP2M DIKTI yang telah mendanai penelitian ini melalui skim Hibah Kompetensi sehingga penelitian ini terlaksana.

Tim Peneliti

DAFTAR ISI

HALAMAN SAMPUL

HALAMAN PENGESAHAN

RINGKASAN

PRAKATA

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB 1. PENDAHULUAN

BAB 2. TINJAUAN PUSTAKA

BAB 3. TUJUAN DAN MANFAAT PENELITIAN

BAB 4. METODE PENELITIAN

BAB 5. HASIL DAN PEMBAHASAN

BAB 6. RENCANA TAHAPAN BERIKUTNYA (untuk laporan tahunan)

BAB 7. KESIMPULAN DAN SARAN

DAFTAR PUSTAKA

LAMPIRAN

- instrumen
- personalia tenaga peneliti beserta kualifikasinya
- HKI dan publikasi

DAFTAR TABEL

Tabel 1. Hasil Analisis Quest Soal Paket A	64
Tabel 2. Hasil Analisis Quest Soal Paket B	67
Tabel 3. Hasil Analisis Quest Soal Paket C	70

DAFTAR GAMBAR

Gambar 1. Roadmap Penelitian	6
Gambar 2. Paradigma Penelitian	16
Gambar 3. Desain Penelitian	18
Gambar 4. Hasil Rekapitulasi Equating Paket A, B, dan C	61
Gambar 5. Grafik rekapitulasi hasil equating Paket A dan B	74
Gambar 5. Grafik rekapitulasi hasil equating Paket B dan C	74
Gambar 5. Grafik rekapitulasi hasil equating Paket A dan C	75
Gambar 8. Kuisioner tentang Kemudahan Akses oleh User (guru)	78

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Ujian nasional (UN) telah memunculkan kontroversi yang berkepanjangan yang masih meninggalkan sejumlah persoalan dan pertanyaan yang menarik untuk dikaji. Kontroversi itu makin mengemuka dengan adanya berbagai permasalahan pada UN tahun 2013 khususnya tingkat SMA/K, diantaranya adalah; penundaan ujian nasional di 11 provinsi, keterlambatan paket soal, kekurangan lembar soal dan lembar jawaban, paket mata pelajaran tertukar, kualitas kertas yang buruk, soal ujian nasional tercecer, tidak bisa mengikuti karena berhadapan dengan hukum, sekolah tidak kebagian soal dan lembar jawaban, materi ujian tak sesuai jadwal, problem UN untuk siswa berkebutuhan khusus, serta pengiriman soal salah daerah.

Permasalahan lain yang muncul, dan dianggap tidak memnuhi rasa keadilan adalah tentang cara penentuan kelulusan dari sistem ujian nasional. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2013 Tentang Kriteria Kelulusan Peserta Didik Dari Satuan Pendidikan dan Penyelenggaraan Ujian Sekolah/Madrasah/Pendidikan Kesetaraan dan Ujian Nasional, salah satu pasalnya mengungkapkan bahwa, nilai akhir (NA) untuk penentuan kelulusan diperoleh dari gabungan nilai sekolah dari mata pelajaran yang diujikan secara nasional dan Nilai UN, yaitu dengan pembobotan 40% Nilai sekolah dari mata pelajaran yang diujikan secara nasional dan 60% dari Nilai UN. Permasalahan yang kemudian muncul berkaitan dengan hal ini adalah belum adanya kesetaraan kualitas assessment yang digunakan untuk penilaian di sekolah (semester 3 (tiga) sampai dengan 5 (lima) untuk tingkat SLTP/SLTA/SMK), sehingga belum dapat menjamin rasa keadilan karena perbedaan kualitas tes yang diberikan.

Permasalahan serupa terjadi juga pada sistem penerimaan mahasiswa baru (SNMPTN), khususnya pada jalur undangan. Keputusan pemerintah menggunakan jalur undangan dan menghapuskan jalur tulis dari Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN) 2013 mengundang banyak kritik, meskipun jalur tulis tetap ada dan dikelola mandiri oleh PTN secara bersama-sama (SBMPTN). Sistem tersebut dinilai berpotensi memicu munculnya upaya manipulasi nilai rapor yang masif di sekolah. Meskipun hal ini telah coba diantisipasi dengan membuat Pangkalan Data Sekolah dan Siswa (PDSS), sebagai basis data yang

berisikan rekam jejak sekolah dan prestasi akademik siswa yang berminat masuk PTN melalui jalur undangan. Kritik yang mencuat adalah penggunaan nilai sekolah yang sampai saat ini belum ada standarisasi tentang kualitas penilaian dan memicu tindak manipulasi pada nilai rapor siswa di sekolah. Sistem ini banyak dikritik karena akan melahirkan bentuk manipulasi masif baru dalam sistem pendidikan di jenjang pendidikan menengah karena kondisinya memang belum siap. Apalagi pada tahun ini, kuota pendaftaran ke PTN melalui jalur undangan disama ratakan padahal kualitas sekolah, kualitas guru yang berdampak pada kualitas soal yang diberikan pada ujian sekolah juga sangat beragam. Kondisi ini berbeda dengan tahun sebelumnya yang diatur berdasarkan status akreditasi masing-masing sekolah, misal sekolah dengan akreditasi A mendapatkan kuota 50 persen dari jumlah siswa untuk mendaftar ke PTN, sedangkan akreditasi B sebanyak 30 persen, lalu akreditasi C sebesar 15 persen.

Karena itulah, hal yang sangat mendesak untuk dicarikan solusinya, adalah menghasilkan sistem assessment yang terstandar melalui penyetaraan kualitas soal ujian sekolah melalui proses equiting dan Bank Soal. Pada kebanyakan program tes dalam skala besar, penyusunan tes-tes yang setara merupakan kegiatan yang sangat penting. Hal ini harus dilakukan untuk penanganan yang cepat apabila terjadi kebocoran tes dan untuk membandingkan hasil tes peserta yang menggunakan tes-tes yang berbeda tersebut. Kegiatan ini dapat dilakukan dengan menggunakan teori responsi butir (*item response theory*). Ada tiga asumsi yang harus dipenuhi dalam teori responsi butir, yaitu: 1) unidimensi, 2) independensi lokal, dan 3) invariansi. Langkah-langkah melakukan kegiatan penyetaraan tes berdasarkan teori responsi butir, yaitu: 1) mengestimasi parameter, 2) mengestimasi skala teori responsi butir dengan transformasi linear, dan 3) menyamakan sekor. Sedangkan rancangan penyetaraan tes yang digunakan menurut teori responsi butir ada tiga, yaitu: 1) rancangan kelompok tunggal, 2) rancangan kelompok equivalen, dan 3) rancangan tes jangkar. Metode-metode yang saat ini dikembangkan untuk melakukan penyetaraan tes menurut teori responsi butir adalah: 1) metode regresi, 2) metode rerata dan sigma (*mean and sigma method*), 3) metode tegar rerata dan sigma (*robust mean and sigma method*), dan 4) metode kurva karakteristik. Berdasarkan kajian dari permasalahan di atas maka dalam enelitian ini, akan dikembangkan model assessment terstandar berbasis CMI (*Computer Management*

Instructional) untuk menjamin kesetaraan kualitas penilaian sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN Jalur Undangan yang berkeadilan.

Bagaimanapun juga, ujian merupakan strategi yang umum digunakan untuk meningkatkan mutu pendidikan manakala sumber daya yang dimiliki oleh suatu negara relatif terbatas. Oleh karena itu, UN memegang peranan strategis di dalam sistem pendidikan di negara berkembang seperti Indonesia. Secara konseptual, ujian merupakan strategi evaluasi yang potensial untuk mendorong peningkatan mutu pendidikan melalui (1) pengendalian mutu lulusan, dan (2) motivator atau pendorong bagi guru, siswa, dan penyelenggara pendidikan dalam meningkatkan upayanya secara optimal. Sedangkan SNMPTN tetap sangat penting sebagai syarat masuk perguruan tinggi merupakan jembatan awal bagi pemerintah dalam hal ini pihak kampus untuk mempersiapkan sumberdaya yang benar-benar berkualitas. Adapun sistem undangan merupakan wujud pengakuan dari pendidikan tinggi atas kinerja dan kredibilitas proses belajar dan hasil evaluasi di jenjang sekolah menengah.

B. Rumusan Masalah

Berdasarkan latar belakang permasalahan di atas maka dapat dirumuskan permasalahan sebagai berikut;

1. Bagaimana mengembangkan tes terstandar sesuai dengan kompetensi inti dari kurikulum 2013 yang dapat digunakan sebagai perangkat equiting (penyetaraan) dalam penyusunan tes di seluruh sekolah sehingga dapat digunakan sebagai penentu kelulusan dalam sistem ujian akhir nasional yang berkeadilan?
2. Bagaimana mengembangkan model assessment terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian di seluruh sekolah sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan?
3. Bagaimana menemukan ukuran sampel minimum, pengaruh panjang tes, panjang tes *anchor* minimum, dan metode enyetaraan tes dalam penyetaraan vertikal model kredit parsial menggunakan *common-item nonequivalent groups design*?
4. Bagaimana menghasilkan basis data untuk digunakan dalam penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan?

C. Luaran Penelitian dan Kontribusi Terhadap Ilmu Pengetahuan

Program penelitian Kompetensi yang secara spesifik ini berkaitan dengan upaya pengembangan model assessment terstandar berbasis CMI (*Computer Management Instructional*) dengan menggunakan metode equiting dan Bank Soal. Penelitian ini sangat penting untuk memberikan kontribusi dalam menyelesaikan permasalahan yang berkembang di masyarakat, khususnya berkaitan dengan UN dan SNMPTN jalur undangan yang mendapat kritikan tajam terutama berkaitan dengan belum terpenuhinya rasa keadilan. Dengan demikian kontribusi hasil penelitian Kompetensi ini yang menghasilkan model yang sekaligus terapannya sangat bermanfaat dalam skala nasional.

Kegiatan dalam penelitian ini dilakukakan menyeluruh mulai dari penyiapan proses penyetaraan vertikal memakai *common-item nonequivalent groups design* dan penentuan koefisien penyetaraan dengan Program *STUIRT*, dan dalam penyetaraan kualitas tes digunakan Program *WinGen2*. Keakuratan metode penyetaraan diukur dengan (a) *Root Mean Square Difference (RMSD)* antara parameter hasil estimasi dan parameter hasil bangkitan; dan (b) *Root Mean Square Error (RMSE)* antara parameter butir hasil estimasi dan parameter seajatnya.. Dengan demikian, diharapkan akan menghasilkan luaran yang bermanfaat baik ditinjau dari segi kepentingan nasional khususnya Kementrian Pendidikan dan Kebudayaan, peningkatan kualitas penilaian, dan peningkatan kualitas peserta didik. Oleh karena itu, luaran dari penelitian ini antara lain adalah:

1. **Model assessment terstandar berbasis CMI (*Computer Management Instructional*) dengan menggunakan metode equiting dan Bank Soal**, untuk penyetaraan kualitas soal dalam pentuan kelulusan UN dan SNMPTN diharapkan mampu memenuhi rasa keadilan peserta tes, sehingga maka model yang dikembangkan dalam penelitian ini dapat dijadikan acuan secara nasional.
2. **Publikasi artikel ilmiah pada jurnal nasional/internasional yang terakreditasi.** Pengembangan Model assessment terstandar berbasis CMI (*Computer Management Instructional*) dengan menggunakan metode equiting dan Bank Soal adalah bersifat aktual dan orisinal karena baru dikembangkan dan belum diteliti secara lebih mendalam, oleh karena itu sangat berpeluang untuk dipublikasikan baik di jurnal nasional maupun internasional.

BAB II

TINJAUAN PUSTAKA

Pada peta jalan penelitian ini, peneliti ingin memberikan gambaran yang jelas tentang status kegiatan penelitian yang berjudul pengembangan model assessment terstandar berbasis CMI (*computer management instructional*) untuk menjamin kesetaraan kualitas penilaian sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan terhadap kemungkinan pengembangan kegiatan tersebut di masa depan dan diakhiri dengan tujuan yang ingin dicapai melalui penelitian ini.

Sejak menyelesaikan Desertasi yang berjudul “Model Pembelajaran Lima Domain Sains dengan Pendekatan Kontekstual untuk Mengembangkan Pembelajaran Bermakna”, ketua tim peneliti telah berkonsentrasi untuk melakukan penelitian pada bidang pengukuran dan evaluasi pendidikan khususnya pengembangan model model pembelajaran atau model implementasi kurikulum di tingkat kelas pembelajaran dan sekaligus evaluasinya. Konsistensi ini ditunjukkan dengan penelitian **Hibah Bersaing (DPPM/Ditlitabmas)** pada tahun 2007-2008 yang berjudul, ” Model KKN-PPL Tematik Pengembangan Kit Praktikum Sains Realistik Hasil *Re-Use* Limbah Anorganik Sebagai Media *Joyfull Learning* untuk Rehabilitasi Pendidikan dan Psikologis di Sekolah Terdampak Erupsi Merapi”, dan dilanjutkan dengan **Hibah Bersaing (DPPM/Ditlitabmas)** tahun 2009 yang berjudul, “Model Kesiapsiagaan Bencana (*Disaster Preparedness*) Dalam Bentuk Pembelajaran Sekolah Darurat Dengan Pendekatan *Fun Learning* Menggunakan Media Pembelajaran Dari Limbah Rumah Tangga Untuk Penanganan Pendidikan di Daerah Pasca Bencana”.

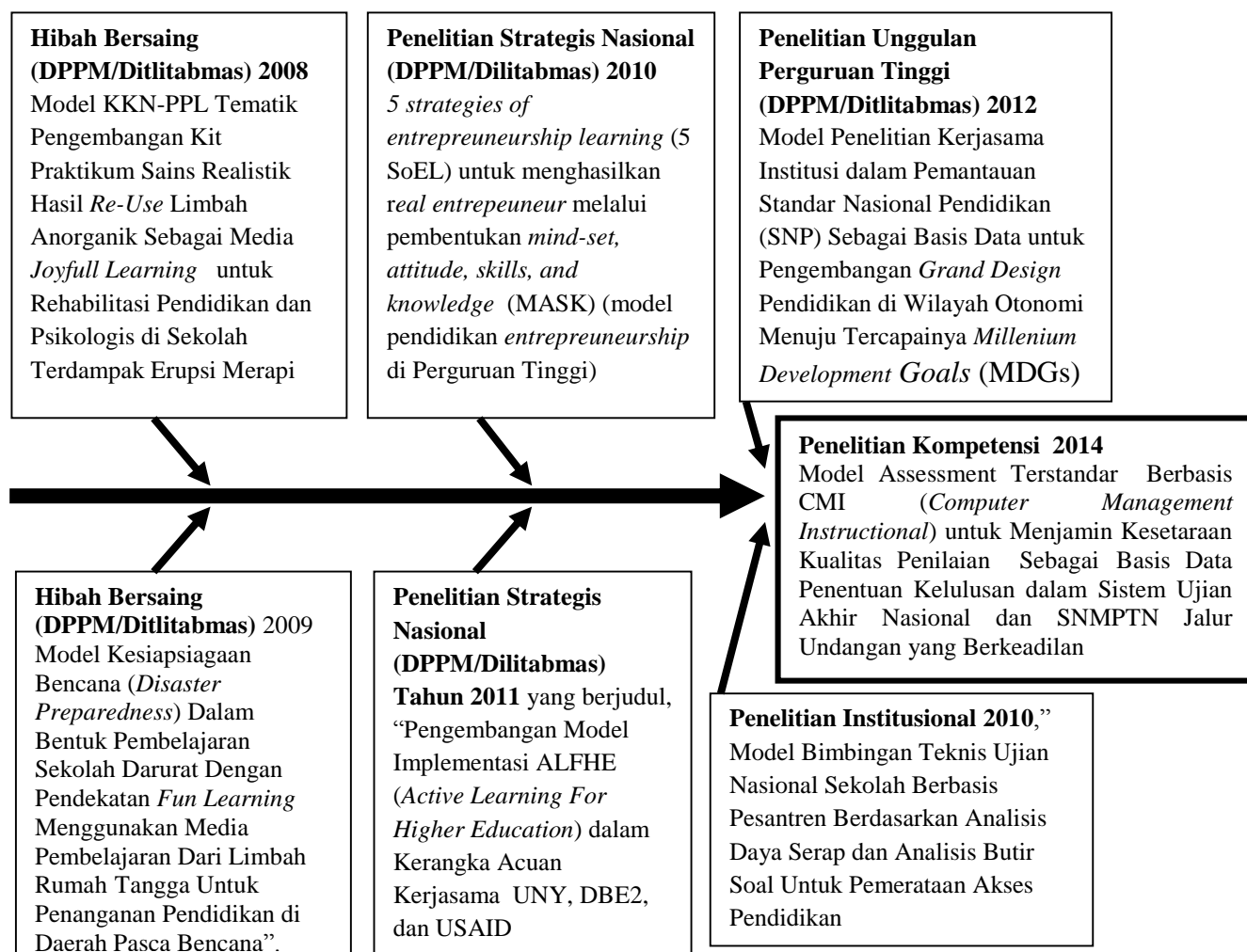
Penelitian tentang model pembelajaran inovatif dan evaluasinya itu dilanjutkan ke jenjang yang lebih tinggi yaitu **Penelitian Strategis Nasional (DPPM/Ditlitabmas) Tahun 2010** yang berjudul, “*5 strategies of entrepreneurship learning (5 SoEL)* untuk menghasilkan *real entrepreneur* melalui pembentukan *mind-set, attitude, skills, and knowledge* (MASK) (model pendidikan *entrepreneurship* di Perguruan Tinggi)“, yang mencoba memberikan alternatif bagi inovasi pengembangan nilai kewirausahaan dalam pembelajaran sains, serta **Penelitian Strategis Nasional (DPPM/Ditlitabmas) Tahun 2011** yang berjudul, “Pengembangan Model Implementasi ALFHE (*Active Learning For Higher Education*) dalam Kerangka Acuan Kerjasama UNY, DBE2, dan USAID”, yang mulai mengembangkan

kerjasama internasional dengan DBE2 USAID dalam mengembangkan model pembelajaran aktif untuk meningkatkan kemampuan bernalar dan penelitian ilmiah peserta didik.

Pada tahun 2011/2012, ketua tim peneliti mencoba mengembangkan lebih lanjut kompetensi penelitian melalui **Penelitian Institusional** yang berjudul, ” Model Bimbingan Teknis Ujian Nasional Sekolah Berbasis Pesantren Berdasarkan Analisis Daya Serap dan Analisis Butir Soal Untuk Pemerataan Akses Pendidikan”, dan sekaligus meningkatkan kerjasama yang lebih luas dengan pemerintah daerah melalui **Penelitian Unggulan Perguruan Tinggi (DPPM/Ditlitabmas)** dengan judul penelitian, ” Model Penelitian Kerjasama Institusi dalam Pemantauan Standar Nasional Pendidikan (SNP) Sebagai Basis Data untuk Pengembangan *Grand Design* Pendidikan di Wilayah Otonomi Menuju Tercapainya *Millenium Development Goals* (MDGs)”. Penelitian kebijakan ini semakin mengukuhkan kompetensi berkaitan dengan pengukuran dalam bidang pendidikan.

Untuk lebih memudahkan dalam memahami jalan peta penelitian ini, maka dapat dilihat dari peta *fish bone* di bawah ini.

Gambar 1. Roadmap Penelitian



Pengembangan kegiatan penelitian ini di masa depan akan lebih terfokus pada pengembangan model implementasi model-model pengukuran, pengembangan assessment dan evaluasinya. Hal ini sangat penting untuk menjamin sukses tidaknya suatu kurikulum diberlakukan sehingga memiliki kemanfaatan dalam skala nasional bahkan lebih.

Polemik yang berkembang di masyarakat bahwa tidak ada assessment terstandar yang layak digunakan untuk penyetaraan kualitas tes dalam penentuan kelulusan ujian nasional dan SNMPTN jalur undangan harusnya dapat diatasi dengan adanya sistem yang baik dan menjamin rasa keadilan bagi seluruh warga negara Indonesia. Pada program pengujian, khususnya dalam skala besar, penyusunan beberapa perangkat tes yang setara adalah salah satu kegiatan penting karena salah satu tugasnya adalah menjaga keamanan perangkat tes tersebut. Pada taraf tertentu kesetaraan beberapa perangkat tes dapat dilaksanakan pada saat mengembangkan tes itu sendiri, tetapi biasanya bervariasi antara perangkat tes yang satu dengan perangkat tes lainnya terutama dalam hal tingkat kesukaran. Hal ini dapat diatasi dengan melakukan kegiatan penyetaraan sekor antar perangkat tes dengan cara yang tepat dan benar. Seringkali dijumpai di sekolah, peserta tes yang berbeda harus diukur dengan tes yang berbeda meskipun tes-tes tersebut belum tentu ekuivalen dan diharapkan dapat mengukur sifat dan tuntutan pencapaian hasil yang dapat dibandingkan (Tumilisar, 2006: 3).

Walaupun sampai taraf tertentu kesetaraan beberapa tes dapat diupayakan pada saat menyusun tes-tes itu sendiri, tetapi pada umumnya variasi taraf sukar antar tes tetap terjadi (Swediati, 1997: 1). Selain itu, penyetaraan tes perlu dilakukan mengingat bahwa menyusun tes yang benar-benar paralel tidaklah mudah. Jadi secara empirik membuat dua tes yang sama, tidak pernah secara sempurna paralel, terandakan atau unidimensional, sehingga sekor-sekor yang dihasilkannya tidak dapat dibandingkan (Gronlund, 1985: 169). Jika hasil tes digunakan untuk menentukan kenaikan kelas atau penjurusan program, tentunya hal itu menjadi tidak adil karena tidak dilakukan ekuivalensi sekor untuk tes yang berbeda tersebut. Oleh karena itu, penting dilakukan penyesuaian sekor-sekor tersebut sehingga peserta tes yang berbeda, menggunakan tes yang berbeda tersebut dapat dibandingkan.

Masalah tersebut dapat diatasi dengan melakukan penyetaraan sekor yang diperoleh dari peserta yang mengambil tes-tes tersebut. Proses statistik dikenal sebagai metode penyetaraan

(*equating*), sudah dikembangkan untuk menangani masalah ini. Dengan kata lain penyetaraan adalah proses untuk menentukan hubungan antara skala sekor dari dua tes atau lebih tes agar sekor-sekor tersebut diperlakukan secara adil. Kegiatan penyetaraan tes dapat dilakukan dengan mengembangkan konversi suatu sistem unit tes ke sistem unit tes yang lain sehingga setelah dikonversi sekor yang berasal dari dua perangkat tes menjadi setara dan dapat dipertukarkan. Kegiatan ini dapat dilakukan dengan menggunakan teori tes kalsik dan teori responsi butir. Dalam artikel ini pembahasan difokuskan pada penerapan teori responsi butir (*item response theory*). Penerapan teori responsi butir dalam penyetaraan tes sangat berguna terutama bagi pengembangan bank soal.

Kebaruan dalam bidang penelitian dan pengabdian kepada masyarakat dari penelitian ini sangat jelas, karena pengembangan model *assessment* terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian sebagai basis data penentuan kelulusan dalam sistem Ujian Akhir Nasional dan SNMPTN jalur undangan yang berkeadilan adalah bersifat aktual dan orisinal karena baru dikembangkan dan belum diteliti secara lebih mendalam, oleh karena itu sangat berpeluang untuk dipublikasikan baik di jurnal nasional maupun internasional.

Pustaka acuan yang actual dan terkini berkaitan dengan *assessment* terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian sebagai basis data penentuan kelulusan dalam sistem Ujian Akhir Nasional dan SNMPTN jalur undangan yang berkeadilan lebih berkaitan dengan penelitian penerapan teori respon butir. Beberapa jurnal yang diacu untuk memperdalam penelitian ini, adalah;

1. Dorans, N. J. (2004). Equating, concordance, and expectation. *Applied Psychological Measurement*, 28 (4),227-246.
2. Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Journal of Educational measurement* (4th ed., pp. 187{220). Westport, CT: Greenwood.
3. Kim, S., von Davier, A. A., & Haberman, S. (2008). Small-sample equating using a synthetic linking function. *Journal of Educational Measurement*, 45, 325{342}.
4. Livingston, S. A., & Kim, S. (2009). The circle-arc method for equating in small samples. *Journal of Educational Measurement*, 46, 330{343}.

5. Lord, F. M. (2009). The standard error of equipercentile equating. *Journal of Educational Statistics*, 7, 165{174}.
6. Skaggs, G. (2005). Accuracy of random groups equating with very small amples. *Journal of Educational Measurement*, 42, 309{330}.

Naga (1992: 24) menyatakan bahwa sekor tes sebagai hasil koreksi dari setiap butir yang dikerjakan peserta tes yang menampilkan jawaban benar atau salah dan membedakan menjadi dua, yaitu: 1) sekor tunggal yaitu sekor satu butir sebagai jawaban dari satu peserta tes; dan 2) sekor komposit yaitu gabungan dari sekor tunggal. Barnard (1996: 16) menyatakan bahwa tidak ada definisi penyetaraan tes yang dapat diterima secara universal. Penyetaraan oleh Peterson, Kolen dan Hoover (1989: 221) didefinisikan sebagai proses yang digunakan untuk memastikan sekor-sekor yang dihasilkan dari pengadministrasian tes-tes dapat digunakan secara bergantian dan Crocker dan Algina (1986: 457), penyetaraan didefinisikan sebagai suatu proses untuk menetapkan sekor-sekor equivalen pada dua instrumen. Penyetaraan sekor adalah suatu prosedur empiris yang diperlukan untuk mentransformasi sekor suatu perangkat tes ke sekor perangkat tes yang lain. Karena merupakan prosedur empiris maka penyetaraan sekor didasarkan pada data sekor tes.

Membuat soal yang setara dalam dua buah paket atau lebih, tentunya tidak mudah atau bahkan tidak mungkin, karena pasti ada perbedaannya. Hal tersebut disebabkan karena hampir tidak mungkin menyusun multi paket tes yang benar-benar paralel (Petersen, Kolen, & Hoover, 1989). Meskipun penyusun tes menggunakan spesifikasi tes yang sama dalam menulis item-itemnya dan hanya merubah angka, tidak ada jaminan bahwa tingkat kesukaran item-item tersebut akan sama. Apalagi kalau yang berbeda adalah kata kunci dan isi dari pilihan jawaban. Menurut Angoff (1971) dan Kolen (1988) seperti yang dikutip dalam Hambleton (1991), metode penyetaraan dibagi menjadi 2 kategori, yaitu: 1) penyetaraan ekipersentil, dan 2) penyetaraan linear (*linear equating*).

Kategori pertamaini merupakan penyempurnaan sekor dengan melakukan perbandingan antara sekor tes X dan Y menjadi ekuivalen apabila urutan persen rangking masing-masing group adalah sama. Selanjutnya untuk penyetaraan skor dalam 2 tes yang berbeda, maka kedua tes tersebut harus diberikan pada kelompok examine yang sama. Selanjutnya pada kategori kedua, diasumsikan bahwa sekor x pada test X dan sekor y pada test Y mempunyai hubungan yang searah/segaris (*linearly related*). Menurut Tumilisar (2006), metode penyetaraan adalah

cara-cara mencari hubungan penyetaraan dua sekor tes dari dua instrumen penelitian yang berbeda dengan menggunakan statistik tertentu, dan pengumpulan datanya dilakukan dengan desain pengumpulan data tertentu. Metode penyetaraan ekipersentil dibagi dua, yaitu:

1. Metode penyetaraan ekipersentil berantai adalah cara mencari ekivalensi ekipersentil dua sekor tes dari dua instrumen penelitian yang berbeda, pengumpulan datanya dilakukan dengan desain nonekivalen tes jangkar dan tes jangkarnya merupakan tes jangkar internal dengan menggunakan statistik tertentu. Ekivalensi ekipersentil dihitung dengan metode penyetaraan ekipersentil langsung secara terpisah pada sekor tes kedua instrumen, masing-masing terhadap tes jangkarnya, tanpa menggunakan populasi sintetik.
2. Metode penyetaraan ekipersentil estimasi frekuensi adalah cara mencari ekivalensi persentil dua sekor tes dari dua instrumen penelitian yang berbeda dengan menggunakan statistik tertentu, dan pengumpulan datanya dilakukan dengan desain nonekivalen tes jangkar dan tes jangkarnya merupakan tes jangkar internal. Ekivalensi ekipersentil dihitung dengan mengestimasi distribusi kumulatif dua sekor tes masing-masing terhadap tes jangkarnya, dengan menggunakan populasi sintetik.

Proses penyetaraan dari beberapa perangkat tes (*equating*) dapat dilakukan dengan dua cara, yaitu penyetaraan secara horizontal dan penyetaraan secara vertikal. Proses penyetaraan yang diperoleh dari dua perangkat tes yang berbeda tetapi mengukur hal yang sama dinamakan penyetaraan horizontal. Adapun proses penyetaraan dari dua kelompok peserta tes yang berbeda dalam jenjang/tingkat pendidikannya, namun diberi perangkat soal yang sama dinamakan penyetaraan vertikal (Crocker & Algina, 1986). Pada dasarnya penyetaraan (*equating*) bertujuan untuk menyetarakan sekor dengan cara membandingkan sekor yang diperoleh dari mengerjakan perangkat tes yang satu dengan sekor yang diperoleh dari mengerjakan perangkat tes lainnya yang dilakukan melalui proses penyetaraan sekor pada kedua perangkat tes tersebut (Hambleton & Swaminthan, 1991). Menurut Zhu (1998), sekor-sekor pada tes A dan tes B dapat disetarakan jika memenuhi empat syarat, yaitu: 1) mengukur kemampuan atau karakteristik yang sama. Sehingga tes-tes yang disusun dari kisi-kisi yang berbeda tidak dapat disetarakan; 2) setelah penyetaraan, distribusi frekuensi sekor pada tes A harus sama seperti distribusi frekuensi sekor pada tes B, sehingga sekor pada tes A dan tes B dapat saling dipertukarkan setelah penyetaraan; 3) penyetaraan tes harus bebas dari data atau pekerjaan peserta tes dalam proses penyetaraan, dan konversi yang berasal dari penyetaraan

harus berlaku bagi semua situasi yang serupa; dan 4) transformasinya harus sama tanpa memperhatikan tes mana yang digunakan sebagai dasar atau referensi konversi, artinya interpretasi skor harus sama baik penyetaraan dari tes A ke tes B atau dari tes B ke tes A.

Lord (1980) mengemukakan gagasan atau ide penyetaraan dalam beberapa implikasi, yaitu: 1) pengukuran tes dengan sifat yang berbeda tidak dapat disetarakan; 2) skor mentah pada tes yang konsisten tidak sama, tidak dapat dilakukan proses penyetaraan; 3) skor mentah pada tes dengan kesukaran yang bervariasi tidak dapat disetarakan karena tes tidak akan konsisten sama pada tingkat kesukaran yang sama; 4) kekeliruan atau kesalahan skor pada tes atau paket A dan B tidak dapat disetarakan kecuali jika kedua tes tersebut benar-benar paralel; dan 5) tes yang sempurna reliabilitasnya dapat dilakukan penyetaraan

Penyetaraan dilakukan dengan cara mengkonversikan satu paket ke paket yang lain, dari paket yang mengukur kemampuan yang sama. Penyetaraan perangkat tes merupakan pembuatan sejumlah keputusan dari skor yang diperoleh dari sebuah paket untuk disesuaikan ke bentuk yang berbeda tingkat kesukarannya. Jika ada paket X lebih sukar dari paket Y, maka penyetaraan paket X ke Y menghasilkan nilai paket X lebih tinggi atau berharga jika disetarakan ke paket Y (Crocker dan Algina, 1986). Ada tiga dasar dalam merancang data untuk diambil atau dianalisis dalam melakukan penyetaraan tes (Kolen & Brennan, 2004), yaitu: 1) desain data yang dikumpulkan dari dua kelompok yang dites pada paket berbeda dengan kisi-kisi sama, dimana pembagian kedua paket tersebut secara acak atau random; 2) untuk proses penyetaraan, salah satu kelompok tes diberikan paket A setelah itu dites kembali dengan paket B, dan satu kelompok lagi diberikan dulu paket B kemudian mengerjakan kembali paket A; dan 3) perbedaan instrumen tes yang diberikan kepada peserta ujian yang berbeda pula. Namun dalam kedua paket tersebut terdapat tes jangkar (*anchor test*) yang diberikan kepada seluruh peserta tes. Tes jangkar itulah yang dijadikan patokan untuk melakukan penyetaraan. Peserta tes dalam hal ini tidak perlu dibagi secara acak atau random walaupun pembagian dengan random juga tidak akan mempengaruhi model ini.

BAB III

TUJUAN DAN MANFAAT PENELITIAN

A. Tujuan Penelitian

Tujuan jangka panjang dari penelitian ini adalah mengembangkan model assessment terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian di seluruh sekolah sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan. Tujuan khusus dari penelitian ini adalah;

1. mengembangkan tes terstandar sesuai dengan kompetensi inti dari kurikulum 2013 yang dapat digunakan sebagai perangkat equiting (penyetaraan) dalam penyusunan tes di seluruh sekolah sehingga dapat digunakan sebagai penentu kelulusan dalam sistem ujian akhir nasional yang berkeadilan,
2. mengembangkan model assessment terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian di seluruh sekolah sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan,
3. menemukan ukuran sampel minimum, pengaruh panjang tes, panjang tes *anchor* minimum, dan metode enyetaraan tes dalam penyetaraan vertikal model kredit parsial menggunakan *common-item nonequivalent groups design*,
4. menghasilkan basis data untuk digunakan dalam penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan.

B. Manfaat Penelitian

Sesuai dengan tujuan dan latar belakang permasalahan, maka manfaat dari penelitian ini. Diantaranya adalah:

1. Manfaat Secara Teoritis

- a. Model assessment terstandar berbasis CMI (*Computer Management Instructional*) dengan menggunakan metode equiting dan Bank Soal merupakan solusi untuk menghasilkan kualitas penilaian yang setara sehingga menjamin kesetaraan kualitas penilaian di seluruh sekolah sebagai basis data penentuan kelulusan dalam sistem ujian

akhir nasional dan SNMPTN jalur undangan yang berdampak pada terpenuhinya rasa keadilan.

- b. Pola pengembangan Model assessment terstandar berbasis CMI (*Computer Management Instructional*) dengan menggunakan metode equiting dan Bank Soal dapat dijadikan referensi yang sangat tepat sebagai best practice penentuan standar penilaian yang dapat diterapkan secara nasional.

2. Manfaat secara praktis

- a. Untuk melakukan mendapatkan hasil penelitian yang dapat menyelesaikan masalah bangsa dan masyarakat dengan fokus bidang pendidikan dalam mengembangkan inovasi sistem penilaian ujian akhir nasional dan seleksi masuk perguruan tinggi yang terstandar dan berkeadilan.
- b. Memberikan peluang yang lebih tinggi bahwa kualitas dan kompetensi dosen peneliti akan lebih baik.
- c. Dapat meningkatkan kualitas sistem penilaian dengan adanya basis data dan bank soal menggunakan teknik penyetaraan soal berbasis CMI dengan cara implementasi di beberapa sekolah yang heterogen sehingga dapat mewakili sekolah di Indonesia sebagai *best practice*
- d. Mendorong perguruan tinggi untuk dapat memanfaatkan fasilitas, dosen, dan laboratorium untuk proses pengembangan inovasi kurikulum pembelajaran sehingga dapat dimanfaatkan untuk penelitian yang dapat berguna bagi negara dan bangsa
- e. Meningkatkan, menguatkan, dan menjaga kesinambungan periset dan institusi untuk melaksanakan Riset di Universitas Negeri Yogyakarta.
- f. Meningkatkan kompetensi periset Universitas Negeri Yogyakarta pada bidang prioritas model penerapan kurikulum dan model model pembelajaran.
- g. Mengembangkan keilmuan terkini dan pemanfaatannya untuk menyelesaikan permasalahan yang berkembang di masyarakat.

BAB IV

METODE PENELITIAN

A. Paradigma dan Disain Penelitian

Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2013 tentang kriteria kelulusan peserta didik dari Satuan Pendidikan dan Penyelenggaraan Ujian Sekolah/Madrasah/Pendidikan Kesetaraan dan Ujian Nasional, menyatakan bahwa Nilai Akhir untuk kelulusan dari Satuan Pendidikan diperoleh dari gabungan Nilai S/M/PK (Nilai Sekolah/Madrasah/Pendidikan Kesetaraan) dari mata pelajaran yang diujikan secara nasional dan Nilai UN, yaitu dengan pembobotan 40% Nilai S/M/PK dari mata pelajaran yang diujikan secara nasional dan 60% dari Nilai UN (Pasal 6, ayat 3).

Ujian Nasional merupakan salah satu penilaian eksternal yang digunakan pemerintah untuk mengumpulkan data pencapaian prestasi belajar peserta didik, sejauh mana prestasi belajar peserta didik mencapai Standar Kompetensi Lulusan (SKL). Di sekolah peserta didik seharusnya sudah terbiasa dengan penilaian hasil belajar yang dilakukan oleh guru sekolah. Sebagaimana diamanatkan oleh Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan Pasal 63 ayat (1): Penilaian pendidikan pada jenjang pendidikan dasar dan menengah terdiri atas: 1) penilaian hasil belajar oleh pendidik; 2) penilaian hasil belajar oleh satuan pendidikan; dan 3) penilaian hasil belajar oleh pemerintah (Wibowo, 2011:9).

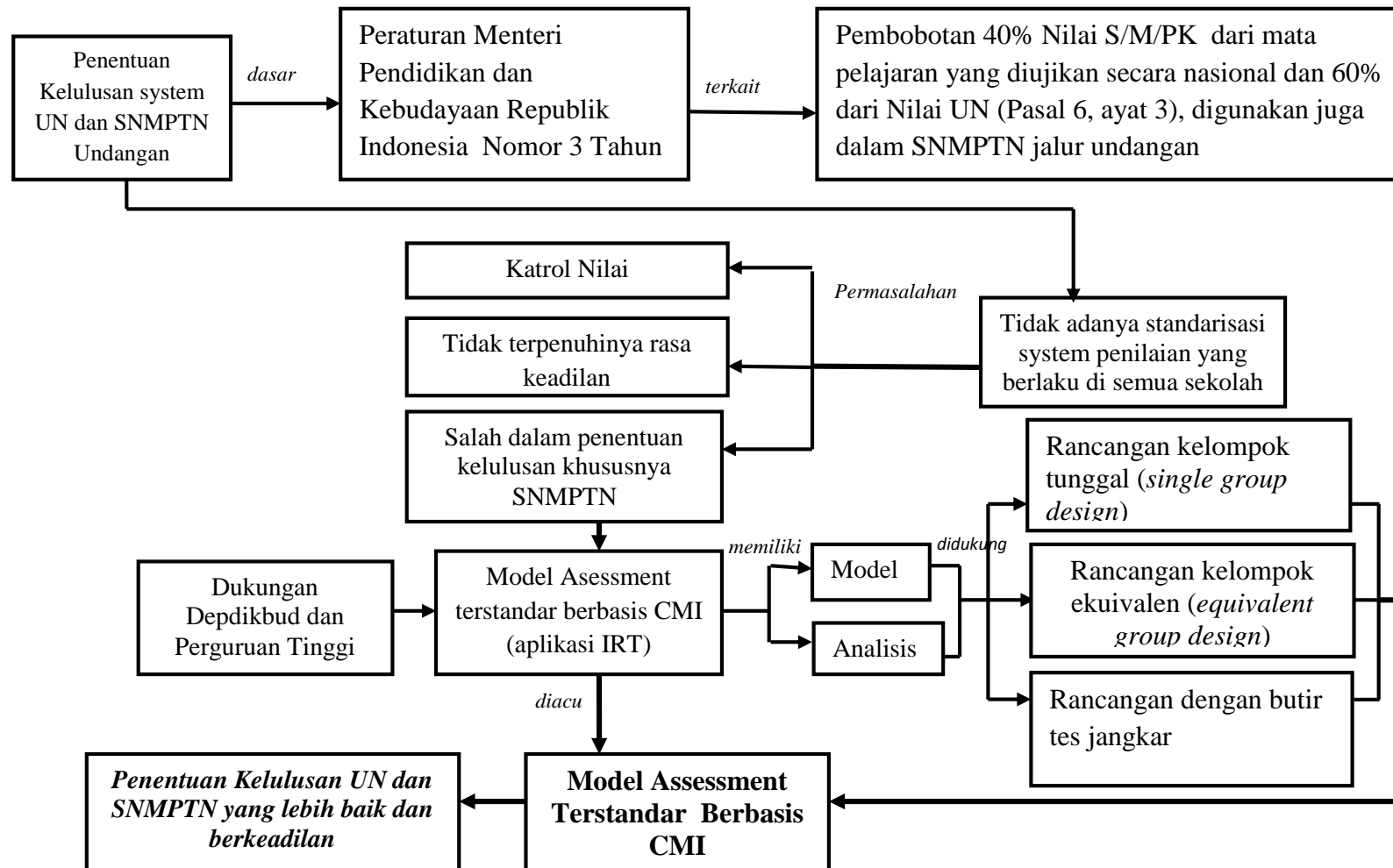
Penilaian hasil belajar oleh pendidik dilakukan secara berkesinambungan untuk memantau proses, kemajuan, dan perbaikan hasil dalam bentuk Ujian Sekolah, ujian tengah semester, ujian akhir semester, dan ujian kenaikan kelas. Penilaian hasil belajar oleh pendidik digunakan untuk menilai pencapaian kompetensi peserta didik; bahan penyusunan laporan hasil belajar; dan memperbaiki proses pembelajaran. Penilaian hasil belajar oleh satuan pendidikan bertujuan menilai pencapaian standar kompetensi lulusan untuk semua mata pelajaran. Penilaian hasil belajar oleh pemerintah dalam bentuk ujian nasional bertujuan untuk menilai pencapaian kompetensi lulusan secara nasional pada mata pelajaran tertentu dalam kelompok mata pelajaran ilmu pengetahuan dan teknologi. Ujian nasional dilakukan secara objektif, berkeadilan, dan akuntabel.

Permasalahan yang kemudian muncul berkaitan dengan hal ini adalah belum adanya kesetaraan kualitas assessment yang digunakan untuk penilaian di sekolah (semester 3 (tiga) sampai dengan 5 (lima) untuk tingkat SLTP/SLTA/SMK), sehingga belum dapat menjamin rasa keadilan karena perbedaan kualitas tes yang diberikan. Hal ini sangat mendesak untuk dicarikan

solusinya, karena nilai sekolah digunakan juga dalam sistem penerimaan mahasiswa baru (SNMPTN) jalur undangan. Karena itulah diperlukan penyetaraan kualitas tes agar terstandar dan memenuhi rasa keadilan. Salah satu cara yang dapat dilakukan adalah dengan mengintegrasikan teori respon butir pada system penilaian sekolah melalui CMI.

Aplikasi teori responsi butir dalam kegiatan penyetaraan tes harus memenuhi dua asumsi dasar yaitu unidimensi dan independensi lokal (Kollen dan Brennan, 2004). Sedangkan langkah-langkah melakukan kegiatan penyetaraan tes menurut teori responsi butir adalah: 1) melakukan estimasi parameter butir dan parameter kemampuan; 2) mengestimasi skala teori responsi butir dengan menggunakan transformasi linear; dan 3) menyamakan sekor, jika menggunakan sekor jawaban yang benar maka dilakukan konversi ke skala jawaban benar dan kemudian ke skala sekor. Oleh karena kegiatan penyetaraan tes memiliki prosedur yang empiris, maka kegiatan ini memerlukan rancangan tertentu yang harus diperhatikan.

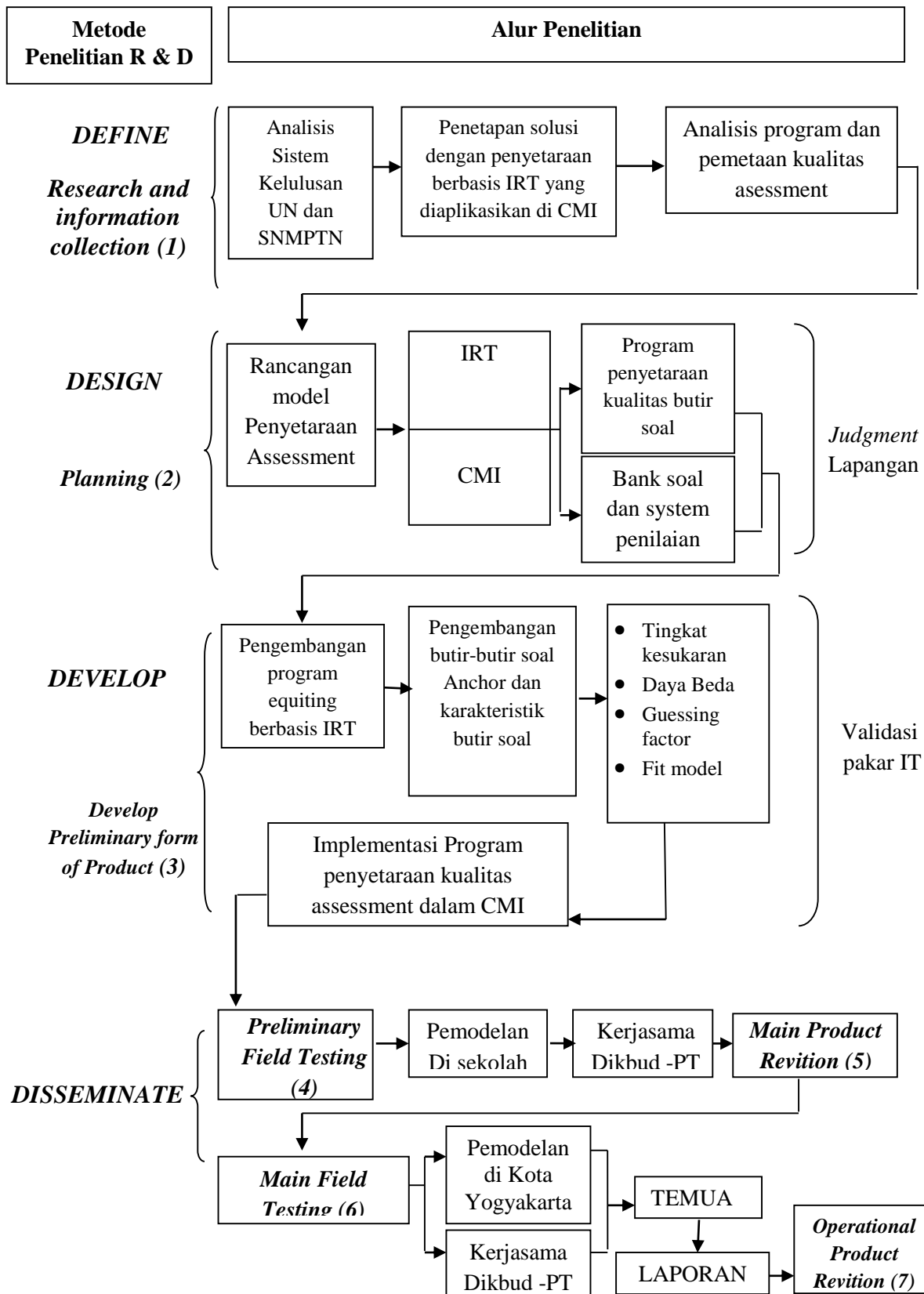
Penyetaraan dilakukan dengan cara mengkonversikan satu paket ke paket yang lain, dari paket yang mengukur kemampuan yang sama. Penyetaraan perangkat tes merupakan pembuatan sejumlah keputusan dari sekor yang diperoleh dari sebuah paket untuk disesuaikan ke bentuk yang berbeda tingkat kesukarannya. Jika ada paket X lebih sukar dari paket Y, maka penyetaraan paket X ke Y menghasilkan nilai paket X lebih tinggi atau berharga jika disetarakan ke paket Y (Crocker dan Algina, 1986). Ada tiga dasar dalam merancang data untuk diambil atau dianalisis dalam melakukan penyetaraan tes (Kolen & Brennan, 2004), yaitu: 1) desain data yang dikumpulkan dari dua kelompok yang dites pada paket berbeda dengan kisi-kisi sama, dimana pembagian kedua paket tersebut secara acak atau random; 2) untuk proses penyetaraan, salah satu kelompok tes diberikan paket A setelah itu dites kembali dengan paket B, dan satu kelompok lagi diberikan dulu paket B kemudian mengerjakan kembali paket A; dan 3) perbedaan instrumen tes yang diberikan kepada peserta ujian yang berbeda pula. Namun dalam kedua paket tersebut terdapat tes jangkar (*anchor test*) yang diberikan kepada seluruh peserta tes. Tes jangkar itulah yang dijadikan patokan untuk melakukan penyetaraan. Peserta tes dalam hal ini tidak perlu dibagi secara acak atau random walaupun pembagian dengan random juga tidak akan mempengaruhi model ini. Paradigma penelitian dijelaskan pada Gambar 1. Perwujudan penyetaraan kualitas assessment berbasis CMI sebagai basis data penentuan kelulusan dan SNMPTN jalur undangan.



Gambar 2. Paradigma Penelitian

Disain penelitian menggunakan metode penelitian *Research and Development* dan alur penelitian yang dijelaskan pada Gambar 2. Fase *define* atau *research and information collection* (Borg dan Gall, 1983: 776) merupakan fase penelitian dan pengumpulan data awal berupa studi literatur, analisis kebutuhan dan studi lapangan. Fase *design* atau *planning* (Borg dan Gall, 1983: 777) merupakan rancangan produk yang akan dihasilkan, meliputi tujuan penggunaan produk, pengguna produk dan deskripsi komponen-komponen produk. Fase *develop* atau *develop preliminary form of product* (Borg dan Gall, 1983: 781) merupakan pengembangan produl awal. Fase *Disseminate* ada empat langkah pengembangan, yaitu *preliminary field testing* (Borg dan Gall, 1983: 782) yang merupakan ujicoba lapangan awal, *main product revision* (Borg dan Gall, 1983: 782) atau revisi hasil ujicoba, *main field testing* (Borg dan Gall, 1983: 783) atau ujicoba lapangan utama serta *operational product revision* (Borg dan Gall, 1983: 784) atau penyempurnaan produk hasil ujicoba lapangan.

Berdasarkan hasil analisis Sistem penentuan kelulusan Ujian Nasional dan Sistem Seleksi Nasional Mahasiswa Perguruan Tinggi Negeri (SNMPTN) jalur undangan pada fase *define*, maka diperoleh daftar permasalahan yang berpotensi untuk menimbulkan rasa ketidakadilan dan ketidaktepatan dalam memprediksi potensi peserta ujian yang lulus dan tidak lulus. Mengacu daftar permasalahan yang ditemukan, maka dilakukan perbaikan sistem dengan menerapkan proses penyetaraan menggunakan Item Response Theory yang diplikasikan dalam sistem Computer Management Instructional (CMI). Desain penelitiannya dapat dilihat pada gambar berikut ini.



Gambar 3. Disain Penelitian

1. Prosedur Penelitian

Mengacu pada desain penelitian tersebut, prosedur penelitian dilaksanakan melalui tahapan-tahapan dalam *research and development* (R & D). Tujuan utama R & D untuk mengembangkan dan memvalidasi suatu program atau model yang akan digunakan agar tujuan menjadi efektif dan siap untuk diimplementasikan. Tahapan-tahapan R & D diformulasikan menjadi model 4-D (*Four-D Models*) (Thiagarajan, 1975: 5) dan disesuaikan Borg dan Gall (1983: 775) yaitu:

a. Define (D-1)/Research and Information Collection

1) Analisis Teori/ Studi Literatur

Tahap ini menganalisis secara teori model model untuk penyetaraan kualitas tes, meliputi:

- Ada tiga jenis rancangan penyetaraan tes yang dapat digunakan, yaitu rancangan kelompok tunggal (*single group design*), rancangan kelompok ekuivalen (*equivalent group design*), dan rancangan dengan butir tes jangkar. Dalam rancangan kelompok tunggal digunakan satu kelompok peserta yang merespons dua perangkat tes (X dan Y). Parameter butir dari kedua perangkat tes diestimasi secara terpisah dengan mengkalibrasi parameter kemampuan peserta tes atau parameter butir. Berdasarkan rancangan ini, dengan mengkalibrasi parameter kemampuan peserta, maka parameter butir dari perangkat tes X dan Y sudah berada pada skala yang sama.
- Idealnya untuk menyetarakan sekor dari beberapa perangkat tes, maka perangkat tes-perangkat tes tersebut diberikan pada responden yang sama. Dengan membandingkan kemampuan peserta tes dari dua/lebih perangkat tes maka penyetaraan dua perangkat tes dapat dilakukan. Kenyataan di lapangan, rancangan ini sulit dilakukan karena adanya faktor kelelahan, belajar, dan adanya faktor latihan untuk tes kedua atau berikutnya. Selain itu, akan terdapat kesulitan dalam hal merencanakan waktu yang cukup bagi responden untuk mengikuti tes lebih dari satu kali (Miyatun dan Mardapi, 2000).
- Pada rancangan kelompok ekuivalen digunakan dua kelompok peserta ekuivalen (K_1 dan K_2) dan dua perangkat tes (X dan Y). Kelompok peserta K_1 mengerjakan perangkat tes X dan kelompok peserta K_2 mengerjakan perangkat tes Y. Mengingat kelompok K_1 dan K_2 adalah ekuivalen, maka kedua kelompok dianggap tunggal. Penentuan konstanta konversi berikutnya seperti rancangan kelompok tunggal. Keuntungan rancangan ini dapat menghindari efek

negatif yang disebabkan karena latihan dan kelelahan peserta tes, sedangkan kekurangannya ada kemungkinan bias yang disebabkan karena tidak mudah untuk membuat distribusi kemampuan dua kelompok peserta tes yang benar-benar ekuivalen (Sukirno, 2007: 310).

- Pada rancangan tes jangkar (*anchor test design*) biasanya digunakan jika masalah keamanan tes menjadi salah satu pertimbangan penting dan memungkinkan untuk menyelenggarakan beberapa tes dalam satu waktu. Pada desain ini masing-masing perangkat tes mempunyai beberapa butir yang sama (*common item*) dan masing-masing kelompok mengerjakan perangkat tes yang berbeda. Pada desain ini terdapat dua variasi, yaitu (Chong dan Osborn, 2005): 1) jika *common item* diperhitungkan dalam pemberian skor disebut *internal common item*; dan 2) jika *common item* tidak diperhitungkan dalam pemberian skor disebut *external common item*.

2) Analisis Tugas/Needs Assesment

Tahap ini mengidentifikasi keterampilan-keterampilan proses utama dan menganalisisnya dalam set-set sub-sub keterampilan yang diperlukan. Analisis ini menjamin kekomprehensifan tugas-tugas dalam penentuan kelulusan sistem ujian nasional dan SNMPTN jalur undangan. Aktivitas penelitian meliputi pengembangan rancangan tes, memuat tujuan penilaian yang akan dilakukan, tempo (waktu yang ditempuh) untuk pelaksanaan pengujian, pesan utama kurikulum (sasaran pembelajaran dan garis besar topik materi uji), indikator butir soal (ciri-ciri penguasaan materi uji dan pencapaian sasaran pembelajaran), serta jumlah dan bentuk butir soal (per-indikator, per topik, dan keseluruhan tes). Sebaran butir soal dalam tes seharusnya memperhatikan keseimbangan tuntutan penguasaan sesuai dengan pesan kurikulum, sehingga memberi nuansa keterwakilan topik bahasan..

3) Analisis Konsep/Needs Assesment

Tahap ini mengidentifikasi konsep-konsep utama yang diajarkan, menyusunnya dalam hirarki dan menguraikan dalam tema-tema utama. Adapun langkah-langkah konstruksi tes yang ditempuh adalah sebagai berikut: 1) menetapkan tujuan tes; 2) analisis kurikulum; 3) analisis buku pelajaran dan sumber materi belajar lainnya; 4) membuat kisi-kisi; 5) penulisan tujuan instruksional khusus; 6) penulisan soal; 7) telaah soal (*face validity*); 8) reproduksi tes terbatas; 9)

uji coba tes; 10) analisis hasil uji coba; 11) revisi soal, dan 12) merakit soal menjadi tes. Metode penyetaraan menurut teori responsi butir berfungsi untuk menentukan konstanta konversi. Hal ini mengingat bahwa penyetaraan antara dua perangkat tes atau lebih dapat dilakukan jika konstanta konversi telah diketahui. Nilai konversi yang dihasilkan kemudian di substitusi dalam persamaan skala pada rancangan penyetaraan yang digunakan. Ada beberapa metode penyetaraan tes yang dapat digunakan dan faktor-faktor yang mempengaruhi keakuratan metode penyetaraan tes. Dalam teori responsi butir terdapat empat metode penyetaraan tes, yaitu: regresi, rerata sigma, rerata dan sigma tegar, dan kurva karakteristik

b. Design (D-2)/Planning

Metode penyetaraan tes yang pertama adalah metode regresi. Penentuan konstanta konversi a dan b menggunakan metode regresi dilakukan dengan memperhatikan respons peserta tes pada kedua perangkat tes X dan Y. Estimasi parameter butir dan parameter kemampuan peserta memenuhi persamaan regresi linier sebagai berikut:

$$y = ax + b + e \text{ dengan } a = r_{xy} S_y/S_x \text{ dan } b = \hat{y} - ax$$

Keterangan:

- y : estimasi kemampuan atau estimasi parameter butir pada perangkat tes Y
- x : estimasi kemampuan atau estimasi parameter butir pada perangkat tes X
- r_{xy} : koefisien korelasi antara X dan Y
- y, x : rerata dari y dan x
- S_y, S_x : simpangan baku dari x dan y
- e : kesalahan dalam penaksiran galat regresi

Penggunaan metode ini bersifat tidak timbal balik (asimetris) sehingga kurang memadai untuk penentuan konstanta konversi apalagi mengingat bahwa penyetaraan dua perangkat tes atau lebih sangat memerlukan syarat invariansi dan timbal balik dari perangkat tes yang disetarakan.

Metode penyetaraan tes yang kedua adalah metode rerata sigma. Pada metode ini, penentuan konstanta konversi α dan β menurut metode rerata dan sigma dilakukan dengan memperhatikan nilai estimasi parameter tingkat kesukaran butir tes pada kedua perangkat tes yaitu b_x dan b_y . Menurut Hambleton & Swaminathan (1985: 26), hubungan antara estimasi parameter butir tes atau parameter

kemampuan peserta pada kedua perangkat tes yang akan disetarakan dan penentuan konstanta konversinya memenuhi persamaan sebagai berikut:

$$y = ax + b \text{ dengan } a = S_y/S_x \text{ dan } b = \hat{y} - ax$$

Metode rerata dan sigma ini bersifat timbal balik sehingga dengan cara yang sama hubungan dari y ke x dapat ditentukan. Namun demikian, menurut Hambleton & Swaminathan (1991: 26) mengemukakan bahwa metode penyetaraan rerata dan sigma ini tidak mempertimbangkan variasi standar error estimasi parameter butir.

Metode penyetaraan tes yang ketiga disebut dengan metode rerata dan sigma tegar. Hambleton dan Swaminathan (1991: 26), menyatakan bahwa dalam metode penyetaraan rerata dan sigma tidak mempertimbangkan variasi estimasi parameter butir. Metode penyetaraan rerata dan sigma tegar mempertimbangkan adanya variasi standard error estimasi parameter butir. Langkah-langkah dalam penentuan konstanta konversi guna penyetaraan tes dengan menggunakan metode ini adalah sebagai berikut (Sukirno, 2007: 312):

1. Penentuan bobot parameter butir (w_i) pada setiap pasangan (b_{xi} dan b_{yi}), yaitu: $w_i = [\max\{v(x_i), v(y_i)\}]^{-1}$ di mana: $i = 1, 2, 3, 4, \dots, k$, $v(x_i)$ dan $v(y_i)$ adalah varian estimasi parameter tingkat kesulitan tes X dan Y.
2. Penentuan penskalaan bobot skala w_i dengan menggunakan rumus: $w_i' = w_i / k$ k = jumlah butir jangkar pada perangkat tes X dan Y.
3. Penghitungan estimasi berbobot tes X dan Y, dengan menggunakan rumus: $x_i' = w_i' x_i$ dan $y_i' = w_i' y_i$
4. Penentuan rerata dan simpangan baku dari estimasi berbobot tes X dan Y, yaitu \bar{x} , \bar{y} , S_x' , S_y' .
5. Penentuan konstanta konversi α dan β dengan menggunakan rerata dan simpangan baku estimasi berbobot dilakukan dengan mensubstitusikan rerata dan simpangan baku bobot estimasi pada persamaan penyamaan skala.

Menurut Stocking dan Lord (Hambleton, 1985) dalam metode penyetaraan rerata dan sigma, proses penentuan konstanta konversi tidak memperhatikan kemungkinan sekor kelompok ekstrim, sedangkan metode penyetaraan rerata dan sigma tegar dapat diperbaiki dengan jalan memperhatikan sekor kelompok ekstrim.

Sedangkan metode keempat yang dapat digunakan dalam penyetaraan tes adalah metode kurva karakteristik. Penentuan konstanta konversi α dan β dengan metode kurva karakteristik, dilakukan dengan memperhatikan nilai estimasi parameter butir tes kedua perangkat soal yaitu x dan y . Metode penyetaraan rerata dan sigma serta metode rerata dan sigma tegar dalam penentuan konstanta konversi hanya memperhitungkan hubungan yang ada antara parameter-parameter kesukaran butir pada perangkat tes yang satu terhadap perangkat tes yang lain. Hubungan antara parameter-parameter daya beda pada kedua perangkat tes belum dipertimbangkan.

Rahayu (2008), menyatakan bahwa metode kurva karakteristik mempertimbangkan informasi dari parameter daya beda butir dan tingkat kesukaran butir dalam penentuan konstanta konversi. Oleh karena itu, dalam metode penyetaraan kurva karakteristik diperhatikan hubungan antara parameter-parameter kesukaran daya beda dan hubungan antara parameter kesukaran butir tes yang akan disetarakan. Selain itu juga dalam metode kurva karakteristik diperhatikan skor asli (*true score*) peserta tes pada kedua perangkat tes.

True Score (t_{xa}) dari peserta tes dengan kemampuan θ_a yang merespon k item dalam perangkat X dan Y adalah:

$$x_a = (\theta_a, b_{xi}, c_{xi}) \quad \text{dan} \quad y_a = (\theta_a, b_{yi}, a_{yi}, c_{yi})$$

Setiap item pada perangkat tes X dan Y memenuhi persamaan:

$$b_{yi} = ab_{xi} + \beta \quad \alpha_{yi} = \alpha \quad c_{yi} = c_{xi} \quad \beta = b_{yi} - ab_{xi}$$

Konstanta α dan β dipilih sedemikian sehingga fungsi F seperti tertera di bawah ini mencapai nilai minimal.

$$F = \sum (x_a - y_i)^2$$

Keterangan :

F : fungsi dari α dan β , yang menunjukkan ketidaksesuaian antara x_a dan y_a

N : jumlah peserta tes

x_a : true score peserta tes pada kemampuan a pada perangkat tes X

y_a : true score peserta tes pada kemampuan a pada perangkat tes Y

Chong dan Osborn (2005) mengemukakan bahwa, ada empat aspek kesetaraan yang harus diperhatikan dalam penyetaraan tes, yaitu: 1. Interferensi. Seberapa jauh skor dari kedua tes dapat digunakan untuk mengukur tujuan yang sama. Misalnya mengukur prestasi akuntansi, mengukur kemampuan berhitung. 2. Konstruk. Seberapa jauh kedua paket tes mengukur konstruk yang sama. 3.

Populasi. Seberapa jauh populasi yang digunakan adalah homogen atau sama. Selain itu faktor-faktor kualitas dan kuantitas yang berhubungan dengan sistem pembelajaran yang harus disetarakan. Artinya sekolah yang memiliki siswa dengan latar belakang sosial ekonomi jauh di bawah, fasilitas sarana prasarana sekolah serba kekurangan, dan guru yang seadanya tidak tepat bila dibandingkan dengan keadaan yang tidak setara. 4. Karakteristik atau kondisi pengukuran. Seberapa jauh kesamaan kondisi pengukuran dilakukan untuk kedua paket tes, baik dari sisi panjang tes, bentuk tes, administrasi tes, waktu tes, tipe soal, dan prosedur tes.

c. *Develop (D-3)/Develop Preliminary form of Product*

Tahap ini memperoleh persetujuan untuk meningkatkan kualitas assessment yang digunakan dalam penentuan kelulusan ujian nasional dan SNMPTN jalur undangan. Sejumlah ahli diminta untuk mengevaluasi program penyetaraan berbasis IRT yang sudah dirancang, meliputi empat metode penyetaraan tes, yaitu: regresi, rerata sigma, rerata dan sigma tegar, dan kurva karakteristik kemudian berbasis *feedback* para ahli model yang telah dihasilkan dimodifikasi/direvisi untuk menjadi lebih tepat, efektif, dan bermanfaat serta teknik kualitasnya tinggi.

d. *Dessiminate (D-4)*

Pengujian perangkat hasil pengembangan (*Preliminary Field Testing*) dilakukan di bekerjasama dengan Dinas Pendidikan dan Kebudayaan dan Perguruan Tinggi Negeri melalui tahap ujicoba dan sudah direvisi (*Main Product Revision*), maka perangkat-perangkat diterapkan dalam penentuan kelulusan pada sistem ujian nasional dan SNMPTN jalur undangan sebenarnya (*Main Field Testing*), kemudian diobservasi segala variabel yang menjadi fokus/tujuan pengembangan.

2. Teknik Evaluasi Program yang Dihasilkan

Metode penyetaraan menurut teori responsi butir berfungsi untuk menentukan konstanta konversi. Hal ini mengingat bahwa penyetaraan antara dua perangkat tes atau lebih dapat dilakukan jika konstanta konversi telah diketahui. Nilai konversi yang dihasilkan kemudian di substitusi dalam persamaan skala pada rancangan penyetaraan yang digunakan. Ada beberapa metode penyetaraan tes yang dapat digunakan dan faktor-faktor yang mempengaruhi keakuratan metode penyetaraan tes. Dalam teori responsi butir terdapat empat metode penyetaraan tes, yaitu: regresi, rerata sigma, rerata dan sigma tegar, dan kurva karakteristik (Angoff, 1982).

Metode penyetaraan tes yang pertama adalah metode regresi. Penentuan konstanta konversi a dan b menggunakan metode regresi dilakukan dengan memperhatikan respons peserta tes pada kedua perangkat tes X dan Y . Estimasi parameter butir dan parameter kemampuan peserta memenuhi persamaan regresi linier sebagai berikut:

$$y = ax + b + e \text{ dengan } a = r_{xy} S_y/S_x \text{ dan } b = \hat{y} - ax$$

Keterangan:

- y : estimasi kemampuan atau estimasi parameter butir pada perangkat tes Y
- x : estimasi kemampuan atau estimasi parameter butir pada perangkat tes X
- r_{xy} : koefisien korelasi antara X dan Y
- y, x : rerata dari y dan x
- S_y, S_x : simpangan baku dari x dan y
- e : kesalahan dalam penaksiran galat regresi

Penggunaan metode ini bersifat tidak timbal balik (asimetris) sehingga kurang memadai untuk penentuan konstanta konversi apalagi mengingat bahwa penyetaraan dua perangkat tes atau lebih sangat memerlukan syarat invariansi dan timbal balik dari perangkat tes yang disetarakan.

Metode penyetaraan tes yang kedua adalah metode rerata sigma. Pada metode ini, penentuan konstanta konversi α dan β menurut metode rerata dan sigma dilakukan dengan memperhatikan nilai estimasi parameter tingkat kesukaran butir tes pada kedua perangkat tes yaitu b_x dan b_y . Menurut Hambleton & Swaminathan (1985: 26), hubungan antara estimasi parameter butir tes atau parameter kemampuan peserta pada kedua perangkat tes yang akan disetarakan dan penentuan konstanta konversinya memenuhi persamaan sebagai berikut:

$$y = ax + b \text{ dengan } a = S_y/S_x \text{ dan } b = \hat{y} - ax$$

Metode rerata dan sigma ini bersifat timbal balik sehingga dengan cara yang sama hubungan dari y ke x dapat ditentukan. Namun demikian, menurut Hambleton & Swaminathan (1991: 26) mengemukakan bahwa metode penyetaraan rerata dan sigma ini tidak mempertimbangkan variasi standar error estimasi parameter butir.

Metode penyetaraan tes yang ketiga disebut dengan metode rerata dan sigma tegar. Hambleton dan Swaminathan (1991: 26), menyatakan bahwa dalam metode penyetaraan rerata dan sigma tidak mempertimbangkan variasi estimasi parameter butir. Metode penyetaraan rerata dan sigma tegar mempertimbangkan adanya variasi standard error estimasi parameter butir. Langkah-langkah dalam

penentuan konstanta konversi guna penyetaraan tes dengan menggunakan metode ini adalah sebagai berikut (Sukirno, 2007: 312):

1. Penentuan bobot parameter butir (w_i) pada setiap pasangan (b_{xi} dan b_{yi}), yaitu:
 $w_i = [\max\{v(x_i), v(y_i)\}]^{-1}$ di mana: $i = 1, 2, 3, 4, \dots, k$, $v(x_i)$ dan $v(y_i)$ adalah varian estimasi parameter tingkat kesulitan tes X dan Y.
2. Penentuan penskalaan bobot skala w_i dengan menggunakan rumus: $w_i' = \frac{1}{k}$ k = jumlah butir jangkar pada perangkat tes X dan Y.
3. Penghitungan estimasi berbobot tes X dan Y, dengan menggunakan rumus: $x_i' = w_i' x_i$ dan $y_i' = w_i' y_i$
4. Penentuan rerata dan simpangan baku dari estimasi berbobot tes X dan Y, yaitu \bar{x} , \bar{y} , S_x' , S_y' .
5. Penentuan konstanta konversi α dan β dengan menggunakan rerata dan simpangan baku estimasi berbobot dilakukan dengan mensubstitusikan rerata dan simpangan baku bobot estimasi pada persamaan penyamaan skala.

Menurut Stocking dan Lord (Hambleton, 1985) dalam metode penyetaraan rerata dan sigma, proses penentuan konstanta konversi tidak memperhatikan kemungkinan sekor kelompok ekstrim, sedangkan metode penyetaraan rerata dan sigma tegar dapat diperbaiki dengan jalan memperhatikan sekor kelompok ekstrim.

Sedangkan metode keempat yang dapat digunakan dalam penyetaraan tes adalah metode kurva karakteristik. Penentuan konstanta konversi α dan β dengan metode kurva karakteristik, dilakukan dengan memperhatikan nilai estimasi parameter butir tes kedua perangkat soal yaitu x dan y . Metode penyetaraan rerata dan sigma serta metode rerata dan sigma tegar dalam penentuan konstanta konversi hanya memperhitungkan hubungan yang ada antara parameter-parameter kesukaran butir pada perangkat tes yang satu terhadap perangkat tes yang lain. Hubungan antara parameter-parameter daya beda pada kedua perangkat tes belum dipertimbangkan.

Rahayu (2008), menyatakan bahwa metode kurva karakteristik mempertimbangkan informasi dari parameter daya beda butir dan tingkat kesukaran butir dalam penentuan konstanta konversi. Oleh karena itu, dalam metode penyetaraan kurva karakteristik diperhatikan hubungan antara parameter-parameter kesukaran daya beda dan hubungan antara parameter kesukaran butir tes yang akan disetarakan. Selain itu juga dalam metode kurva karakteristik diperhatikan sekor asli (*true score*) peserta tes pada kedua perangkat tes.

True Score (t_{xa}) dari peserta tes dengan kemampuan θ_a yang merespon k item dalam perangkat X dan Y adalah:

$$x_a = (\theta_a, b_{xi}, c_{xi}) \quad \text{dan} \quad y_a = (\theta_a, b_{yi}, a_{yi}, c_{yi})$$

Setiap item pada perangkat tes X dan Y memenuhi persamaan:

$$b_{yi} = \alpha b_{xi} + \beta \quad \alpha_{yi} = \alpha \quad c_{yi} = c_{xi} \quad \beta = b_{yi} - \alpha b_{xi}$$

Konstanta α dan β dipilih sedemikian sehingga fungsi F seperti tertera di bawah ini mencapai nilai minimal.

$$F = (x_a - y_i)^2$$

Keterangan :

F : fungsi dari α dan β , yang menunjukkan ketidaksesuaian antara x_a dan y_a

N : jumlah peserta tes

x_a : true score peserta tes pada kemampuan a pada perangkat tes X

y_a : true score peserta tes pada kemampuan a pada perangkat tes Y

Chong dan Osborn (2005) mengemukakan bahwa, ada empat aspek kesetaraan yang harus diperhatikan dalam penyetaraan tes, yaitu: 1. Interferensi. Seberapa jauh sekor dari kedua tes dapat digunakan untuk mengukur tujuan yang sama. Misalnya mengukur prestasi akuntansi, mengukur kemampuan berhitung. 2. Konstruk. Seberapa jauh kedua paket tes mengukur konstruk yang sama. 3. Populasi. Seberapa jauh populasi yang digunakan adalah homogen atau sama. Selain itu faktor-faktor kualitas dan kuantitas yang berhubungan dengan sistem pembelajaran yang harus disetarakan. Artinya sekolah yang memiliki siswa dengan latar belakang sosial ekonomi jauh di bawah, fasilitas sarana prasarana sekolah serba kekurangan, dan guru yang seadanya tidak tepat bila dibandingkan dengan keadaan yang tidak setara. 4. Karakteristik atau kondisi pengukuran. Seberapa jauh kesamaan kondisi pengukuran dilakukan untuk kedua paket tes, baik dari sisi panjang tes, bentuk tes, administrasi tes, waktu tes, tipe soal, dan prosedur tes.

3. Subyek Penelitian

Dalam langkah ujicoba lapangan awal dan ujicoba terbatas model assessment terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan ini digunakan subyek penelitian adalah siswa SMP/MTs dan SMA/K di beberapa sekolah yang ada di 5 Kabupaten Kota di Provinsi DIY dengan mengambil sampel

yang heterogen mulai dari sekolah unggulan, menengah dan tertinggal serta secara geografis dari wilayah Kota kabupaten, Kota Kecamatan dan daerah yang jauh dari akses perkotaan (Rural and urban).

BAB V

HASIL YANG DICAPAI

Penelitian yang telah dilakukan pada tahun kedua hibah kompetensi ini adalah; (1) Software CMI Sistem Informasi Penilaian SMA (SIPSMA), (2) Manual user book untuk program CMI SIPSMA, (3) *Proceeding of International Conference On Research, Implementation And Education Of Mathematics And Sciences (ICRIEMS) 2015, Yogyakarta State University, 17-19 May 2015* (ISBN : 978-979-96880-8-8, lampiran 1), (4) *Proceeding of International Conference On Educational Research and Innovation (ICERI), Institute of Research and Community Services Yogyakarta State University, 6-7 May 2015* (ISSN: 2443-1753, dan publikasi jurnal nasional terakreditasi/ jurnal internasional; (1) *Acceptance* Jurnal Kependidikan (LPPM UNY), Jurnal Nasional Terakreditasi (2) proses review Jurnal Ilmu Pendidikan (Universitas Negeri Malang), Jurnal Nasional terakreditasi dan (3) Submit jurnal internasional di International Journal of Education and Learning (Canadian Centre of Science and Education) terindeks Scopus

Penelitian pada tahap kedua ini, dari rencana tiga tahun penelitian, telah sesuai dengan perencanaan awal yaitu; (1) mengembangkan model assessment terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian di seluruh sekolah sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan, (2) menemukan ukuran sampel minimum, pengaruh panjang tes, panjang tes *anchor* minimum, dan metode enyetaraan tes dalam penyetaraan vertikal model kredit parsial menggunakan *common-item nonequivalent groups design*, dan , (3) menghasilkan Bank Soal untuk Ujian Akhir Sekolah yang digunakan dalam penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan.

A. DESKRIPSI SIPSMA BERBASIS CMI

Aplikasi CMI-SIPSMA merupakan suatu sistem yang berbasis *client-server* dimana mesin komputer *client* hanya diintegrasikan dengan sistem *end-user* CMI-SIPSMA dan *client-requirements*. Sedangkan mesin komputer *server* dapat diintegrasikan dengan sistem basis data (*database*) dan *server-requirements*. Aplikasi CMI-SIPSMA juga dapat diaplikasikan pada satu mesin komputer aja yang memiliki keseluruhan bagian sistem yaitu sistem *end-user* CMI-SIPSMA, *server* dan *client-requirements*,

beserta sistem basis data (*database*).

Pada aplikasi CMI-SIPSMA, keamanan dan hak aksesnya dikembangkan dengan tingkat keamanan Pengguna (*User*) dan Peran Pengguna (*User Role*). Setiap Pengguna didasarkan atas tiap-tiap individu guru pada tiap sekolah. Hanya pengguna “admin” yang bertindak sebagai Super User, peran pengguna Administrator dan sebagai pengguna “default” dengan tidak didasarkan atas individu guru.

Tipe Peran Pengguna dan hak aksesnya berupa :

1. ADMINISTRATOR

Peran pengguna Administrator hampir memiliki hak akses penuh terhadap semua menu, fungsi, dan modul yang ada dalam aplikasi. Sebagian besar wewenang dan tugas dari Administrator berupa :

- a. Buka dan tutup tahun ajaran
- b. Pengelolaan profil sekolah
- c. Pengelolaan pengguna
- d. Pengelolaan guru
- e. Pengelolaan peserta didik
- f. Pengelolaan kelas
- g. Pengelolaan mata pelajaran
- h. Pengelolaan ekstra kurikuler
- i. Pengelolaan kelompok mata pelajaran
- j. Pengelolaan pengajaran
- k. Pengelolaan pindah sekolah

2. GURU WALI KELAS

Peran pengguna Guru Wali Kelas mempunyai hak akses berupa wewenang individu guru yang juga bertindak sebagai wali kelas. Sebagian besar wewenang dan tugas dari Guru Wali Kelas berupa :

- a. Pengelolaan pindah sekolah
- b. Keseluruhan penilaian
- c. Laporan Hasil Belajar

3. GURU MATA PELAJARAN

Peran pengguna Guru Mata Pelajaran hak akses berupa wewenang individu guru yang hanya bertindak sebagai guru mata pelajaran. Sebagian besar wewenang dan tugas dari Guru Mata Pelajaran berupa :

- a. Penilaian harian
- b. Penilaian aspek pengetahuan
- c. Penilaian aspek keterampilan

d. Penilaian aspek sikap spiritual dan sosial

Aplikasi yang dirancang dan telah di ujicobakan adalah CMI (*Computer Management Instruction*) pengembangan sistem penilaian untuk sekolah menengah atas (SMA) selanjutnya disingkat dengan istilah Sistem Informasi dan Penilaian Sekolah Menengah Atas (SIPSMA), dengan konfigurasi sebagai berikut:

1. Petunjuk Instalasi SIPSMA Berbasis CMI

1.1. Spesifikasi Minimum

Spesifikasi minimum perangkat keras (*hardware*) untuk menjalankan aplikasi CMI-SIPSMA beserta fitur-fiturnya adalah :

1. Mesin komputer dengan Processor Intel Core atau AMD dan RAM 2 GB.
2. Printer.

Spesifikasi minimum perangkat lunak (*software*) untuk menjalankan aplikasi CMI-SIPSMA beserta fitur-fiturnya adalah :

1. Sistem operasi Windows XP Professional Service Pack 3 (*client*).
2. Sistem operasi Windows Server 2003 Standard Service Pack 2 (*server*).
3. Microsoft .NET Framework 3.5 atau Microsoft .NET Framework 4.5.1.
4. Microsoft SQL Server 2008 Express dengan Microsoft SQL Server Management Studio.
5. Crystal Report 2008 Runtime.
6. Microsoft Office Word dan Excel.
7. Adobe Reader.
8. Driver printer.

1 Instalasi dan Konfigurasi

1.2. Instalasi

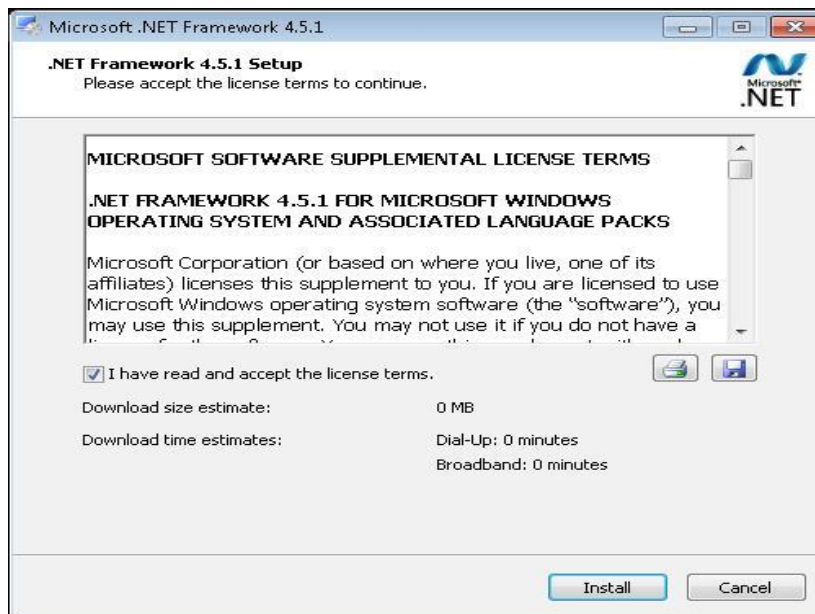
1.2.1. Pra Instalasi

Sebelum memulai keseluruhan instalasi, pastikan Windows user account yang digunakan untuk menginstal merupakan user dengan tipe **Administrator**. Pastikan juga pada konfigurasi Windows Control Panel - > Regional and Language, tab Formats --> Format merupakan format **English (United States)** dan tab Administrative --> Current Language for non-Unicode programs merupakan format **English (United States)**.

1.2.2. Instalasi Microsoft .NET Framework 3.5 atau 4.5.1

Microsoft .NET Framework 3.5 diinstal pada mesin komputer *server* dan *client*.. Instalasi Microsoft .NET Framework 3.5 untuk sistem operasi Windows XP menggunakan file **dotnetfx35.exe** dan 4.5.1 untuk sistem operasi Windows XP, Windows Vista, Windows 7 atau Windows 8 menggunakan file **NDP451-KB2858728-x86-x64-AllOS-ENU.exe**. Jika di komputer sudah pernah terinstall Microsoft .NET Framework maka tidak perlu lagi diinstal selama tidak ada masalah. Klik 2 kali pada file tersebut untuk memulai proses penginstalan.

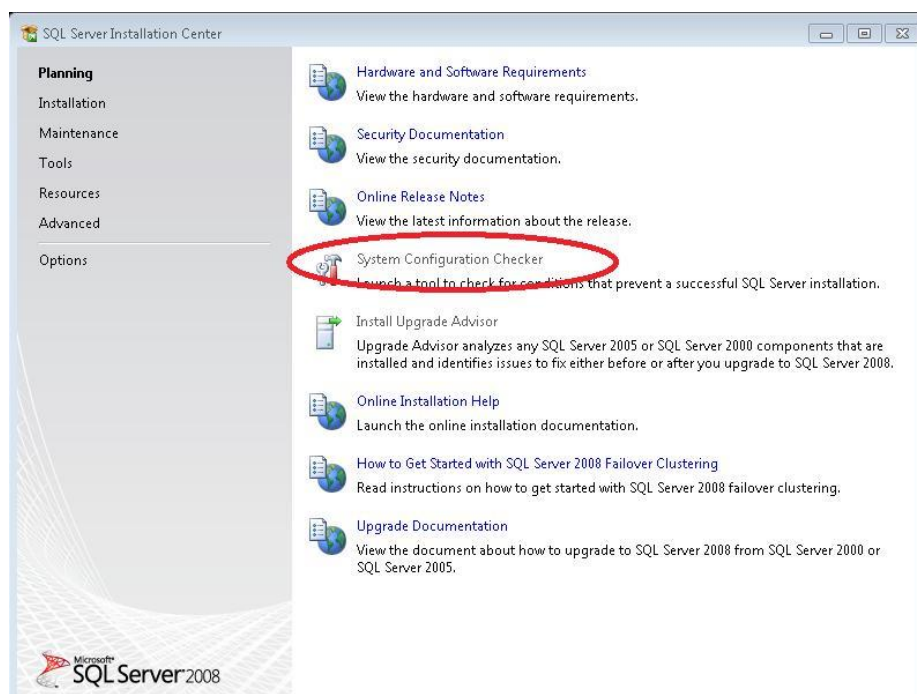
Centang pada pilihan “I have read and accept the license terms”. Klik tombol Install untuk memulai proses penginstalan.



1.2.3. Instalasi Microsoft SQL Server 2008 Express With Tool

Instalasi Microsoft SQL Server 2008 Express With Tool menggunakan file **SQLEXPRT_x86_ENU.exe** untuk komputer dengan arsitektur x86 dan **SQLEXPRT_x64_ENU.exe** untuk komputer dengan arsitektur x64. Klik 2 kali pada file tersebut untuk memulai proses penginstalan.

Klik pada “System Configuration Checker” untuk memeriksa konfigurasi sistem komputer pada saat pra-instalasi.



Apabila masih terdapat laporan status yang mempunyai status “Failed”, instalasi tidak dapat dilanjutkan. Oleh karena itu permasalahan harus segera diatasi dahulu sebelum melanjutkan proses instalasi. Klik OK untuk lanjut ke tahap berikutnya.

Pada tahap berikut ini, mesin komputer *server* dan *client* memiliki proses dan konfigurasi yang berbeda.

Untuk mesin komputer *server*, klik Select All untuk mencentang keseluruhan pilihan, sehingga keseluruhan pilihan harus dipilih seperti Database Engine Services, SQL Server Replication, Management Tools-Basic, dan SQL Client Connectivity. Sedangkan untuk mesin komputer *client*, hanya centang pilihan dari Management Tools-Basic dan SQL Client Connectivity saja.

Klik Next untuk lanjut ke tahap berikutnya.

Crystal Report 2008 Runtime hanya diinstal pada mesin komputer *client*. Instalasi Crystal Report 2008 Runtime menggunakan file **CRRedist2008_x86.msi** untuk komputer dengan arsitektur x86 dan **CRRedist2008_x64.msi** untuk komputer dengan arsitektur x64. Klik 2 kali pada file CRRedist2008_x86.msi untuk memulai proses penginstalan.

1.2.5. Instalasi CMI-SIPSMA

CMI-SIPSMA hanya diinstal pada mesin komputer *client*. Instalasi CMI-SIPSMA menggunakan file **setup.exe** atau dengan file **SetupCMI-SIPSMA.msi**. Klik 2 kali pada file setup.exe untuk memulai proses penginstalan.

Pastikan folder instalasi berada pad “C:\andihermanto\CMI-SIPSMA\”. Untuk aplikasi CMI-SIPSMA yang hanya digunakan untuk satu pengguna Windows account, centang Just Me. Untuk aplikasi CMI-SIPSMA yang digunakan untuk lebih dari satu pengguna Windows account, centang Everyone. Klik Next untuk lanjut ke tahap berikutnya.

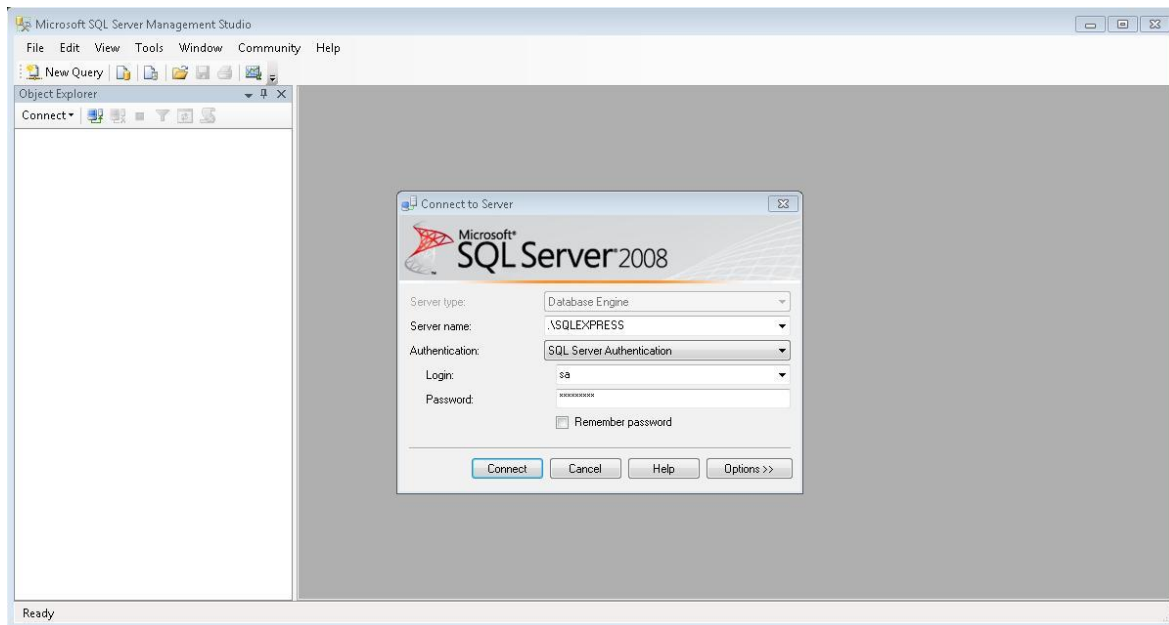


1.3. Konfigurasi

1.3.1. Konfigurasi Database

Petunjuk konfigurasi ini hanya digunakan untuk sistem *server* dan *client* yang berada pada satu mesin komputer. Untuk konfigurasi sistem dengan sistem *server* dan *client* yang berada pada mesin komputer yang berbeda, silahkan konsultasikan terlebih dahulu ke Pengembang Aplikasi CMI-SIP SMA.

Untuk memulai konfigurasi *database*, klik 2 kali *shortcut* SQL Server Management Studio pada Start Menu Windows --> Microsoft SQL Server 2008. Akan muncul jendela utama Microsoft SQL Server Management Studio dengan jendela Connect to Server. Apabila jendela Connect to Server tidak muncul, klik Connect pada tab Object Explorer, kemudian klik Database Engine.



Isi Server name dengan “.SQLEXPRESS”. Pilih Authentication dengan SQL Server Authentication. Isi Login dengan “sa”. Isi Password dengan password yang telah dimasukkan sebelumnya pada tahap instalasi Microsoft SQL Server 2008 Express. Klik tombol Connect pada jendela Connect to Server.

Setelah login Connect to Server berhasil, akan muncul deskripsi SQL Server Database Engine pada tab Object Explorer dengan folder-foldernya. Klik kanan pada folder Database. Klik pada New Database.

Akan muncul jendela New Database. Isikan Database name dengan “CMI-SIPSMA”. Klik OK untuk lanjut ke tahap berikutnya.

Pada halaman utama, klik menu File, pilih Open, kemudian klik File. Cari file dengan ekstensi .sql **DB_CMI-SIPSMA.sql**. Pilih file tersebut dan klik Open.

Selanjutnya akan muncul jendela DB_CMI-SIPSMA.sql pada halaman utama. Klik Execute pada toolbar. Jika berhasil maka akan keluar status Query executed successfully pada footer tab Messages. Tutup Microsoft SQL Server Management Studio setelah proses konfigurasi selesai.


1.3.2. Konfigurasi Sistem CMI-SIPSMA

Petunjuk konfigurasi ini hanya digunakan untuk sistem *server* dan *client* yang berada pada satu mesin komputer. Untuk konfigurasi sistem dengan sistem *server* dan *client* yang berada pada mesin komputer yang berbeda, silahkan konsultasikan terlebih dahulu ke Pengembang Aplikasi CMI-SIPSMA.

Pastikan pada konfigurasi Windows Control Panel - > Regional and Language, tab Formats --> Format merupakan format **English (United States)** dan tab Administrative --> Current Language for non-Unicode programs merupakan format **English (United States)**.

Untuk memulai konfigurasi CMI-SIPSMA yang telah terinstal, klik 2 kali shortcut CMI-SIPSMA pada Start Menu -> CMI-SIPSMA atau klik 2 kali shortcut CMI-SIPSMA pada desktop. Akan muncul aplikasi CMI-SIPSMA dengan halaman Utama dan jendela Dialog dengan pesan “Kesalahan : The ConnectionString property has not been initialized”. Klik OK untuk lanjut ke tahap berikutnya.

Setelah itu akan muncul jendela Konfigurasi Server. Pilih KONEKSI LAN (Autentikasi SQL Server), kemudian isikan Server Name dengan “.SQLEXPRESS”. Login dengan “sa”. Isi Password dengan password yang telah dimasukkan sebelumnya pada tahap instalasi Microsoft SQL Server 2008

Express. Klik tombol Simpan  maka akan muncul halaman Log In CMI-SIPSMA dan aplikasi CMI-SIPSMA telah siap digunakan.

1.4. Uninstalasi

Untuk uninstal aplikasi CMI-SIPSMA, pilih dan klik 2 kali CMI-SIPSMA pada daftar program di Windows Control Panel - > Programs -> Program and Features (Uninstall or change a program). Klik tombol Yes untuk mengkonfirmasi.

2. Penggunaan

2.1. Memulai Aplikasi CMI-SIPSMA

Untuk memulai aplikasi, klik 2 kali shortcut CMI-SIPSMA pada Start Menu -> CMI-SIPSMA atau klik 2 kali shortcut CMI-SIPSMA pada desktop. Apabila tahapan konfigurasi benar, maka akan muncul aplikasi CMI-SIPSMA dengan halaman Utama dan halaman Login.


Pada saat memulai dan login aplikasi CMI-SIPSMA untuk pertama kalinya, hanya ada satu pengguna yaitu “admin” yang bertindak sebagai pengguna “default”, peran pengguna Administrator dan merupakan Super User. Pengguna “admin” mempunyai detail pengguna sebagai berikut :

User Name : admin
Password : admin

Penting untuk diingat, setelah pengguna “admin” telah login, diharapkan untuk segera **mengganti password** pengguna “admin” melalui menu Sistem --> Ganti Password atau pada Sistem --> Pengelolaan Pengguna.

2.2. Login


Pada halaman Login, masukkan data Username dan Password.

- Setelah keseluruhan data telah sesuai, klik tombol Masuk  untuk login.
- Setelah login berhasil maka halaman Utama CMI-SIPSMA beserta menu akan dapat diakses.

2.3. Tahun Ajaran


2.3.1. Buka Tahun Ajaran

Halaman Buka Tahun Ajaran digunakan untuk memulai tahun ajaran baru sekolah.

- Masukkan data Tahun Ajaran dan pilih Semester yang sesuai, kemudian klik tombol Buka Tahun Ajaran .
- Setelah itu terlihat status Tahun Ajaran dan Semester pada footer halaman Utama akan berubah sesuai dengan periode ajaran yang telah dimasukkan.




2.3.2. Tutup Tahun Ajaran

Halaman Tutup Tahun Ajaran digunakan untuk menutup tahun ajaran sekolah.

- Pada Tahun Ajaran dan Semester akan muncul periode ajaran yang sedang aktif.
- Klik tombol Tutup Tahun Ajaran , kemudian terlihat status Tahun Ajaran dan Semester pada footer halaman Utama akan berubah.


2.4. Profil Sekolah




Pada halaman Profil Sekolah, masukkan data NPSN/NSS, Nama Sekolah, Alamat Sekolah, Kode Pos Sekolah, Telepon Sekolah, Kelurahan, Kecamatan, Kabupaten/Kota, Provinsi, Website Sekolah dan Email Sekolah.


- Untuk memasukkan data Kepala Sekolah, klik tombol Pilih , akan muncul halaman Daftar Guru.
- Pilihlah salah satu data guru yang sesuai dan yang akan menjadi kepala sekolah dengan mengklik tombol Pilih .
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.
- Untuk mengelola data guru pada halaman Daftar Guru, masuklah ke halaman Pengelolaan Guru pada menu Pengelolaan pada halaman Utama.




2.5. Pengelolaan Pengguna

Halaman Pengelolaan Pengguna dan tabel data pengguna akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.


Klik tombol Tambah  pada toolbar halaman Pengelolaan Pengguna untuk menambah data pengguna aplikasi. Maka akan muncul halaman Pengelolaan Detail Pengguna.

- Masukkan data User Name, Password, Password Ulang dan pilih Peran Pengguna.
- Untuk memasukkan data Nama Guru, klik tombol Pilih , akan muncul halaman Daftar Guru.
- Pilihlah salah satu data guru yang sesuai dan yang akan menjadi pemilik akun pengguna dengan mengklik tombol Pilih .
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.

Klik tombol Ubah/Tampil  pada toolbar halaman Pengelolaan Pengguna untuk mengubah data pengguna aplikasi. Maka akan muncul halaman Pengelolaan Detail Pengguna.

- Sesuaikan data User Name, Password, Password Ulang dan pilih Peran Pengguna.
- Untuk menyesuaikan data Nama Guru, klik tombol Pilih , akan muncul halaman Daftar Guru.
- Pilihlah salah satu data guru yang sesuai dan yang akan menjadi pemilik akun pengguna dengan mengklik tombol Pilih .
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan perubahan data.

2.6. Ganti Password

Pada halaman Ganti Password, masukkan dan sesuaikan data Password dan Password Ulang. Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan perubahan data.

2.7. Ganti Pengguna


Digunakan untuk logout dari aplikasi atau digunakan untuk mengganti akun pengguna.


2.8. Keluar


Digunakan untuk keluar dari aplikasi dan mengakhiri penggunaan aplikasi. Klik tombol Yes untuk konfirmasi keluar aplikasi.


2.9. Pengelolaan Guru

Halaman Pengelolaan Guru dan tabel data guru akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.

Klik tombol Tambah  pada toolbar halaman Pengelolaan Guru untuk menambah data guru. Maka akan muncul halaman Pengelolaan Detail Guru.


- Masukkan data Nama Guru, NIP dan NUPTK.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.


Klik tombol Ubah/Tampil  pada toolbar halaman Pengelolaan Guru untuk mengubah data guru. Maka akan muncul halaman Pengelolaan Detail Guru.

- Sesuaikan data Nama Guru, NIP dan NUPTK.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan perubahan data.

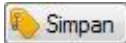
2.10. Pengelolaan Peserta Didik

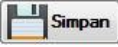
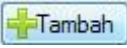
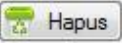

Halaman Pengelolaan Peserta Didik dan tabel data peserta didik akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.

Klik tombol Tambah  pada toolbar halaman Pengelolaan Peserta Didik untuk menambah data peserta didik. Maka akan muncul halaman Pengelolaan Detail Peserta Didik.

- Masukkan data Nama Peserta Didik, Nomor Induk Siswa Nasional, Tempat Tanggal Lahir, pilih data Jenis Kelamin, pilih data Agama, masukkan data Status Dalam Keluarga, Anak ke, Alamat Peserta Didik, Nomor Telepon Rumah, Sekolah Asal, Diterima di Sekolah ini Di Kelas dan Pada Tanggal.
- Untuk memasukkan data Foto Peserta Didik, klik tombol Pilih . Pilihlah salah satu foto yang sesuai pada Open Dialog Explorer dengan mengklik tombol Open.
- Masukkan data Nama Ayah, Nama Ibu, Alamat Orang Tua, Nomor Telepon Rumah, Pekerjaan Ayah, Pekerjaan Ibu, Nama Wali Peserta Didik, Alamat Wali Peserta Didik, Nomor Telepon

Rumah, Pekerjaan Wali Peserta Didik. Pilih data Peminatan kemudian klik tombol Simpan



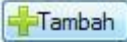
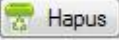




- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data Peserta Didik.
- Menambah mata pelajaran pada tabel data mata pelajaran lintas minat menggunakan tombol Tambah  pada halaman Pengelolaan Detail Peserta Didik, dan akan menampilkan Daftar Mata Pelajaran. Pilihlah mata pelajaran yang akan dimasukkan ke dalam data mata pelajaran lintas minat yang telah ditentukan.
- Menghapus mata pelajaran pada tabel data mata pelajaran lintas minat dengan menggunakan tombol Hapus  pada halaman Pengelolaan Detail Peserta Didik.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data mata pelajaran lintas minat.

2.11. Pengelolaan Kelas

Halaman Pengelolaan Kelas, tabel data kelas dan tabel data peserta didik akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.

Klik tombol Tambah  pada toolbar halaman Pengelolaan Kelas untuk menambah data kelas. Maka akan muncul halaman Pengelolaan Detail Kelas.

- Pilih data Tingkat Kelas, masukkan data Nama Kelas.
- Untuk memasukkan data Wali Kelas, klik tombol Pilih , akan muncul halaman Daftar Guru. Pilihlah salah satu data guru yang sesuai dan yang akan menjadi wali kelas dengan mengklik tombol Pilih .
- Menambah peserta didik pada tabel data peserta didik menggunakan tombol Tambah  pada halaman Pengelolaan Detail Kelas, dan akan menampilkan Daftar Peserta Didik. Pilihlah peserta didik yang akan dimasukkan ke dalam kelas yang telah ditentukan.
- Menghapus peserta didik pada tabel data peserta didik dengan menggunakan tombol Hapus  pada halaman Pengelolaan Detail Kelas.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.


Klik tombol Hapus  pada toolbar halaman Pengelolaan Kelas untuk menghapus salah satu data Kelas.


- Pilih terlebih dahulu data Kelas pada tabel data kelas.
- Klik tombol Hapus pada toolbar halaman Pengelolaan Kelas.
- Klik tombol Yes untuk konfirmasi penghapusan data.


2.12. Pengelolaan Mata Pelajaran


Halaman Pengelolaan Mata Pelajaran dan tabel data mata pelajaran akan muncul bersamaan

dengan aktifnya tombol-tombol pada toolbar.

Klik tombol Tambah  pada toolbar halaman Pengelolaan Mata Pelajaran untuk menambah data mata pelajaran. Maka akan muncul halaman Pengelolaan Detail Mata Pelajaran.


- Masukkan data Nama Mata Pelajaran dan Kelompok Mata Pelajaran.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.


Klik tombol Ubah/Tampil  pada toolbar halaman Pengelolaan Mata Pelajaran untuk mengubah data mata pelajaran. Maka akan muncul halaman Pengelolaan Detail Mata Pelajaran.

- Sesuaikan data Nama Mata Pelajaran dan Kelompok Mata Pelajaran.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan perubahan data.

2.13. Pengelolaan Ekstra Kurikuler

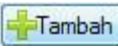



Halaman Pengelolaan Ekstra Kurikuler dan tabel data ekstra kurikuler akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.

Klik tombol Tambah  pada toolbar halaman Pengelolaan Ekstra Kurikuler untuk menambah data ekstra kurikuler. Maka akan muncul halaman Pengelolaan Detail Ekstra Kurikuler.

- Masukkan data Nama Ekstra Kurikuler.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.

2.14. Pengelolaan Kelompok Mata Pelajaran


Halaman Pengelolaan Kelompok Mata Pelajaran dan tabel data mata pelajaran yang telah dikelompokkan akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar. Pilih terlebih dahulu tingkat kelas yang akan diubah pengelompokan mata pelajarannya.

- Menambah daftar mata pelajaran pada tabel kelompok mata pelajaran menggunakan tombol Tambah  pada halaman Pengelolaan Kelompok Mata Pelajaran, dan akan menampilkan Pengelolaan Detail Mata Pelajaran. Pilihlah mata pelajaran dengan tombol Pilih , akan muncul halaman Daftar Mata Pelajaran. Pilihlah salah satu data mata pelajaran yang sesuai dengan mengklik tombol Pilih . Masukkan data KKM yang telah ditentukan. Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.

2.15. Pengelolaan Pengajaran

Halaman Pengelolaan Pengajaran, tabel data guru dan tabel data pengajaran akan muncul


bersamaan dengan aktifnya tombol-tombol pada toolbar.

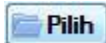

Pilih salah satu guru yang akan dikelola dan penunjukan pada kelas dan mata pelajaran. Klik tombol Ubah/Tampil  pada toolbar halaman Pengelolaan Pengajaran untuk mengubah data pengajaran. Maka akan muncul halaman Pengelolaan Pengajaran – Ubah.

2.16. Pengelolaan Pindah Sekolah

2.16.1. Pindah Masuk


Halaman Pengelolaan Pindah Sekolah (Masuk) dan tabel data pindah masuk akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.



Klik tombol Tambah  pada toolbar halaman Pengelolaan Pindah Sekolah (Masuk) untuk menambah data pindah sekolah. Maka akan muncul halaman Detail Pindah Masuk.

- Pilih data Nama Peserta Didik dengan klik tombol Pilih .
- Setelah Peserta Didik telah dipilih, masukkan Tahun Ajaran dan pilih data Semester.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.

2.16.2. Pindah Keluar


Halaman Pengelolaan Pindah Sekolah (Keluar) dan tabel data pindah keluar akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.



Klik tombol Tambah  pada toolbar halaman Pengelolaan Pindah Sekolah (Keluar) untuk menambah data pindah sekolah. Maka akan muncul halaman Detail Pindah Masuk.

- Pilih data Nama Peserta Didik dengan klik tombol Pilih .
- Setelah Peserta Didik telah dipilih, masukkan Tahun Ajaran dan pilih data Semester.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.

2.17. Equating

Halaman Equating dan tabel data paket akan muncul bersamaan dengan aktifnya tombol-tombol pada toolbar.

Klik tombol Tambah  pada toolbar halaman Equating untuk menambah data paket soal. Maka akan muncul halaman Detail Equating.

- Masukkan data Nama Paket.
- Untuk memasukkan data File Item, klik tombol Pilih . Pilihlah salah satu file .txt yang sesuai pada Open Dialog Explorer dengan mengklik tombol Open. Apabila file .txt yang dipilih sesuai dengan kriteria makan data item akan muncul pada tabel data.
- Setelah keseluruhan data telah sesuai, klik tombol Simpan  untuk menyimpan data.

2.18. Analisis Butir (Excel)

Digunakan untuk membuka file “ANALISIS BUTIR SOAL DENGAN MS.EXCEL.xls”

2.19. Iteman

2.19.1. Jalankan Iteman

Digunakan untuk menjalankan ataupun membuka program Iteman.

2.19.2. Pedoman

Digunakan untuk membuka file “PEDOMAN ANALISIS BUTIR TES.docx”

2.20. Quest

2.20.1. Jalankan Quest

Digunakan untuk menjalankan ataupun membuka program QUEST.

2.20.2. Panduan Praktik

Digunakan untuk membuka file “Panduan Praktik QUEST.docx”

2.20.3. Teori dan Cara Pembacaan

Digunakan untuk membuka file “Teori dan Cara Membaca Hasil Analisi Butir Program QUEST.pdf”

2.21. Notepad

Digunakan untuk menjalankan ataupun membuka program Notepad.

2.22. Jalankan Program

Digunakan untuk menjalankan ataupun membuka file aplikasi.

2.23. Tentang CMI-SIPSMA

Digunakan untuk menampilkan informasi singkat tentang aplikasi CMI-SIPSMA.

B. ANALISIS DATA HASIL DESEMINASI

Wacana yang berkembang di masyarakat bahwa Sistem Seleksi Masuk Perguruan Tinggi Negeri (SMPTN) jalur undangan yang menggunakan nilai raport sekolah dan piagam prestasi dianggap kurang adil karena penerapan standar soal yang berbeda untuk setiap sekolah di Indonesia. Perbedaan

ini terlihat dari adanya variasi tingkat kesukaran soal yang tidak pernah disetarakan antara satu sekolah dengan sekolah lainnya. Variasi kualitas soal ini semakin mencolok ketika berbeda daerah dan wilayah, misalnya di Indonesia Barat dan Indonesia Timur, sekolah di wilayah perkotaan dan sekolah di wilayah pedesaan apalagi pedalaman, sekolah swasta dan negeri, sekolah unggulan dan sekolah tertinggal. Karena dalam kurikulum telah disertai silabus dan indikator pencapaian (Standar Kompetensi/Kompetensi inti (SK/KD), Kompetensi Dasar (KD) dan Standar Kompetensi Lulusan (SKL), maka seharusnya ada standar kualitas soal tertentu yang harus diberikan pada siswa untuk mengukur pencapaian kompetensinya, seharusnya kualitas terutama tingkat kesukaran soal yang digunakan di sekolah tersebut sama atau setara. Paket-paket soal yang digunakan secara empiris disetarakan skornya dengan menggunakan metode Test Score Equating. Penyetaraan skor adalah suatu prosedur empiris yang diperlukan untuk mentransformasi skor suatu perangkat tes ke skor perangkat tes yang lain. Karena merupakan prosedur empiris maka penyetaraan skor didasarkan pada data skor tes (Weiss, 1983).

Membuat soal yang setara dalam dua buah paket atau lebih, tentunya tidak mudah atau bahkan tidak mungkin, karena pasti ada perbedaannya. Hal tersebut disebabkan karena hampir tidak mungkin menyusun multi paket tes yang benar-benar parallel (Petersen, Kolen, & Hoover, 1989). Meskipun penyusun tes menggunakan spesifikasi tes yang sama dalam menulis item-itemnya dan hanya merubah angka, tidak ada jaminan bahwa tingkat kesukaran item-item tersebut akan sama (Kolen, 1984). Apalagi kalau yang berbeda adalah kata kunci dan isi dari pilihan jawaban.

Menurut Angoff (1971) dan Kolen (1988) seperti yang dikutip dalam Hambleton (1991), metode Equating ini dibagikan kepada 2 kategori; yaitu (1) *Equipercentile equating*, dan (2) *Linear Equating*.

Kategori pertama ini merupakan penyempurnaan skor dengan melakukan perbandingan antara skor tes X dan Y menjadi ekuivalen apabila urutan persen rangking masing-masing group adalah sama. Selanjutnya untuk penyetaraan skor dalam 2 tes yang berbeda, maka kedua tes tersebut harus diberikan pada kelompok examinee yang sama. Selanjutnya pada kategori kedua, diasumsikan bahwa skor x pada test X dan skor y pada test Y mempunyai hubungan yang searah/segaris (*linearly related*).

Proses penyetaraan dari beberapa perangkat tes (equating) dapat dilakukan dengan dua cara, yaitu penyetaraan secara horizontal dan penyetaraan secara vertikal. Proses penyetaraan yang diperoleh dari dua perangkat tes yang berbeda tetapi mengukur hal yang sama dinamakan penyetaraan horizontal. Adapun proses penyetaraan dari dua kelompok peserta tes yang berbeda dalam jenjang/tingkat

pendidikannya, namun diberi perangkat soal yang sama dinamakan penyetaraan vertikal (Crocker & Algina, 1986). Pada dasarnya equating bertujuan untuk menyetarakan skor dengan cara membandingkan skor yang diperoleh dari mengerjakan perangkat tes yang satu dengan skor yang diperoleh dari mengerjakan perangkat tes lainnya yang dilakukan melalui proses penyetaraan sekor pada kedua perangkat tes tersebut (Hambleton & Swaminthan, 1990).

Menurut Zhu (1998), skor-skor pada tes A dan tes B dapat disetarakan jika memenuhi empat syarat:

1. Mengukur kemampuan atau karakteristik yang sama. Sehingga tes-tes yang disusun dari kisi-kisi yang berbeda tidak dapat disetarakan.
2. Setelah equating, distribusi frekuensi skor pada tes A harus sama seperti distribusi frekuensi skor pada tes B, sehingga skor pada tes A dan tes B dapat saling dipertukarkan setelah equating.
3. Equating tes harus bebas dari data atau pekerjaan peserta tes dalam proses equating, dan konversi yang berasal dari equating harus berlaku bagi semua situasi yang serupa.
4. Transformasinya harus sama tanpa memperhatikan tes mana yang digunakan sebagai dasar atau referensi konversi, artinya interpretasi skor harus sama baik equating dari tes A ke tes B atau dari tes B ke tes A.

Dalam buku *Fundamental of Item Response Theory* (Hambleton et al, 1991) Lord (1980) mengemukakan gagasan atau ide equating dalam beberapa implikasi, diantaranya;

1. Pengukuran tes dengan sifat yang berbeda tidak dapat di setarakan atau diequating.
2. Skor mentah pada tes yang konsisten tidak sama, tidak dapat diproses equating.

Skor mentah pada tes dengan kesukaran yang bervariasi tidak dapat disetarakan karena tes tidak akan konsisten sama pada tingkat kesukaran yang sama. Fallible (kekeliruan/kesalahan) Skor pada tes atau paket Y dan X tidak dapat disetarakan kecuali jika kedua test tersebut benar-benar paralel.

Tes yang sempurna reliabilitasnya dapat dilakukan equating

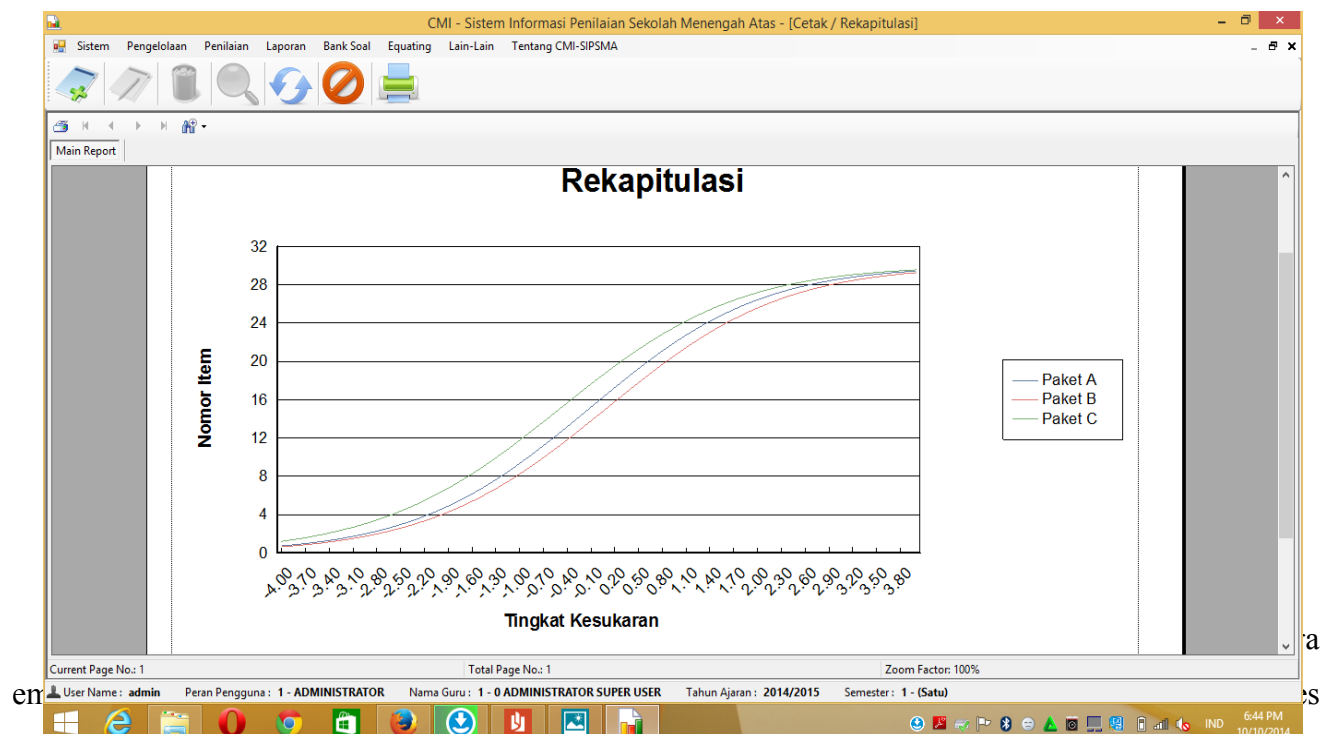
Penyetaraan dilakukan dengan cara mengkonversikan satu paket ke paket yang lain, dari paket yang mengukur kemampuan yang sama. Penyetaraan perangkat tes merupakan pembuatan sejumlah keputusan dari skor yang diperoleh dari sebuah paket untuk disesuaikan ke bentuk yang berbeda tingkat kesukarannya. Jika ada paket X lebih sukar dari paket Y, maka penyetaraan paket X ke Y menghasilkan nilai paket X lebih tinggi atau berharga jika disetarakan ke paket Y (Crocker dan Algina, 1986).

Ada tiga dasar dalam merancang data untuk diambil atau dianalisis dengan equating (Crocker dan Algina, 1986), (Yi, Kim dan Brennan, 2007), yaitu;

1. Design data yang dikumpulkan dari dua kelompok atau group yang di tes paket berbeda dengan kisi-kisi sama, dimana pembagian kedua paket tersebut secara acak atau random.
2. Untuk proses penyetaraan, salah satu kelompok tes diberikan paket A setelah itu di tes kembali dengan paket B, dan satu kelompok lagi diberikan dulu paket B kemudian mengerjakan kembali paket A.
3. Perbedaan instrumen tes yang diberikan kepada peserta ujian yang berbeda pula. Namun dalam kedua paket tersebut terdapat cammon item atau anchor test yang diberikan kepada seluruh peserta tes. Anchor itulah yang dijadikan patokan untuk melakukan equating. Peserta tes dalam hal ini tidak perlu dibagi secara acak atau random walaupun pembagian dengan random juga tidak akan mempengaruhi moel ini. (Crocker dan Algina, 1986).

Ilustrasi dari ketiga rancangan equating dari uraian di atas, dapat dilihat seperti yang tampak pada tabel berikut ini;

Gambar 4. Hasil Rekapitulasi Equating Paket A, B, dan C



equating yang benar, maka memungkinkan konversi secara langsung hasil-hasil ujian peserta tes yang mengambil paket yang berbeda. Tujuan dari equating adalah menyetarakan skor dari A ke B atau dari

B ke A pada kisi-kisi yang sama, karena dua paket tersebut tidak dapat langsung dibandingkan jika belum di setarakan. Penyetaraan tersebut dilakukan melalui rumus transpormasi.

$$Y' = a(X - c) + d$$

Sebagai ilustrasi dalam review ini, peneliti mencoba mengemukakan proses equating antara soal Ujian Sekolah yang diberikan di 16 SMA Indonesia sebagai sekolah mitra dalam deseminasi, yaitu:

1. Paket A: 6 Sekolah di Nusa Tenggara Barat (SMA 1 Labuhan Haji, SMA 1 Sukamulia, SMA 1 Sakra, SMA 1 Keruak, SMA 1 Terara, SMA 1 Pringsela)
2. Paket B: 2 sekolah di Lampung (SMAN 3 lampung dan SMAN 2 Lampung)
3. Paket C : 2 sekolah di Kalimantan tengah (SMAN 4 Palangkaraya dan SMAN 2 Palangkaraya)
4. Paket D: 4 Sekolah di Jawa Barat (SMAN 2 Majalengka, SMAN 1 Ciparay Kabupaten Bandung, SMAN 2 Banjar, dan SMAN Talagasari Kabupaten Karawang)

Instrumen tes adalah bentuk pilihan ganda (5 alternatif jawaban), yang terdiri dari 9 paket dengan 10 soal item anchor (item bersama); Satu paket digunakan untuk Ujian Sekolah matapelajaran Fisika di SMAN se Lombok Timur, 1 paket di Lampung, 1 Paket di Kalimantan Tengah, 2 paket di DIY dan 4 paket di Jawa Barat. Masing-masing paket memiliki panjang tes 40 item (10 item anchor yang disiapkan oleh peneliti dan 30 soal yang dibuat oleh MGMP atau guru masing-masing). Responden adalah siswa SMA mitra yang mengikuti ujian sekolah untuk materi pelajaran Fisika. Tingkat kesukaran soal diolah dengan software QUEST, lalu dilakukan Equating menggunakan metode linear equating random groups design.

Adapun prosedur pelaksanaan equating secara linear equating random groups design melalui persamaan seperti yang dikemukakan oleh Croker dan Algina (1986) yaitu;

1. Jika nilai X dijadikan sebagai patokan (format) pertama, maka persamaannya adalah

$$Y' = a(X - c) + d$$

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

2. Jika nilai Y dijadikan sebagai patokan (format) pertama, maka persamaannya adalah

$$X' = a(Y - c) + d$$

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_y \quad \text{serta } d = \mu_x$$

Dari analisis dengan teori respon butir dengan program QUEST, maka diperoleh informasi statistik terhadap ketiga Paket soal (proses equating dilakukan terhadap hasil analisis Quest yang dikelompokkan menjadi 4 wilayah; yaitu Paket Jawa barat, Paket Lampung, Paket Kalimantan Tengah dan Paket Nusa Tenggara Barat). Hasil analisis butir tes menggunakan Quest yang dijadikan input untuk dilakukan equating menggunakan program equating di Computer Management System yang dikembangkan peneliti dapat dilihat di bawah ini!:

Tabel 1. Contoh Hasil Analisis Quest Butir Soal Ujian Sekolah

UJIAN SEKOLAH: FISIKA
SMAN Selong dan Labuhan Haji

Item Estimates (Thresholds) In input Order 7/ 5/15 5:42
all on all (N = 167 L = 40 Probability Level= .50)

ITEM NAME		[SCORE MAXSCR]		THRSH	INFT	OUTFT	INFT	OUTFT
			1		MNSQ	MNSQ	t	t
1	item 1		111 167		.08	1.13	1.37	1.9 2.8
					.17			
2	item 2		146 167		-1.25	1.10	1.20	.6 .9
					.24			
3	item 3		68 165		1.19	1.01	1.01	.2 .1
					.17			
4	item 4		149 167		-1.43	.98	.96	.0 -.1
					.26			
5	item 5		78 163		.92	1.02	1.05	.4 .5
					.16			
6	item 6		85 167		.77	1.04	1.03	1.0 .4
					.16			

7	item 7		129	167		-.50		.98	1.03	-.1	.2
						.19					
8	item 8		137	167		-.81		.94	.84	-.5	-.8
						.21					
9	item 9		147	166		-1.36		.90	.82	-.5	-.6
						.25					
10	item 10		143	166		-1.13		1.01	1.11	.1	.6
						.23					
11	item 11		98	166		.42		1.10	1.12	1.9	1.2
						.16					
12	item 12		138	166		-.89		.95	.96	-.3	-.1
						.21					
13	item 13		67	165		1.22		.97	1.02	-.6	.2
						.17					
14	item 14		150	167		-1.50		.85	.62	-.8	-1.5
						.26					
15	item 15		94	166		.53		1.13	1.20	2.8	1.8
						.16					
16	item 16		137	166		-.84		.84	.73	-1.2	-1.5
						.21					
17	item 17		141	166		-1.03		.94	1.18	-.4	.9
						.23					
18	item 18		36	164		2.17		1.02	1.05	.2	.4
						.20					

SMAN Selong dan Labuhan Haji

Item Estimates (Thresholds) In input Order

7/ 5/15 5:42

all on all (N = 167 L = 40 Probability Level= .50)

ITEM NAME	SCORE MAXSCR	THRSH	INFT	OUTFT	INFT	OUTFT
	1	MNSQ	MNSQ	t	t	

19 item 19	130 166	-.56	.91	.86	-.8	-.9
------------	---------	------	-----	-----	-----	-----

	.20
--	-----

--	--

20 item 20	143 167	-1.09	.87	.75	-.9	-1.2
------------	---------	-------	-----	-----	-----	------

	.23
--	-----

--	--

21 item 21	148 167	-1.37	.86	.68	-.7	-1.3
------------	---------	-------	-----	-----	-----	------

	.25
--	-----

--	--

22 item 22	53 167	1.63	.95	1.06	-.6	.5
------------	--------	------	-----	------	-----	----

	.17
--	-----

--	--

23 item 23	36 167	2.18	.95	1.33	-.4	1.8
------------	--------	------	-----	------	-----	-----

	.19
--	-----

--	--

24 item 24	158 167	-2.20	1.03	1.12	.2	.4
------------	---------	-------	------	------	----	----

	.35
--	-----

--	--

25 item 25	141 167	-.99	1.01	1.14	.1	.7
------------	---------	------	------	------	----	----

	.22
--	-----

--	--

26 item 26	109 167	.13	.96	.93	-.6	-.6
------------	---------	-----	-----	-----	-----	-----

	.17
--	-----

--	--

27 item 27	114 167	-.01	1.11	1.21	1.5	1.7
------------	---------	------	------	------	-----	-----

	.17
--	-----

--	--

28 item 28	108 165	.13	1.04	1.01	.6	.2
------------	---------	-----	------	------	----	----

	.17
--	-----

29 item 29		46 164		1.83		1.10 1.33 1.2 2.2
				.18		
30 item 30		73 166		1.07		1.09 1.17 2.0 1.6
				.16		
31 item 31		70 167		1.16		.97 .96 -.6 -.4
				.16		
32 item 32		96 167		.49		1.07 1.13 1.5 1.3
				.16		
33 item 33		41 167		2.01		.91 1.09 -1.0 .7
				.19		
34 item 34		76 166		.99		1.00 1.04 .0 .5
				.16		
35 item 35		93 167		.56		.93 .92 -1.5 -.8
				.16		
36 item 36		40 166		2.04		.92 .84 -.8 -1.1
				.19		
37 item 37		132 167		-.61		.96 1.11 -.3 .7
				.20		
38 item 38		106 167		.22		.96 .91 -.7 -.8
				.17		
39 item 39		159 167		-2.33		.95 .69 .0 -.7
				.37		
40 item 40		154 167		-1.80		.98 .87 .0 -.3
				.30		

Mean				.00		.99 1.01 .1 .2
SD				1.27		.08 .18 1.0 1.1

Tabel 2. Contoh Hasil Analisis Quest Untuk Item Ujian Sekolah

SMAN Selong dan Labuhan Haji

Case Estimates In input Order

7/ 5/15 5:43

all on all (N = 167 L = 40 Probability Level= .50)

NAME	SCORE MAXSCR	ESTIMATE	ERROR	INFIT	OUTFT	INFT	OUTFT
		MNSQ	MNSQ	t	t		
1 001	25 40	.68	.38	1.45	1.53	2.27	1.52
2 002	25 40	.68	.38	1.41	1.55	2.10	1.58
3 003	30 40	1.45	.41	1.32	1.71	1.46	1.43
4 004	28 40	1.13	.40	.87	.79	-.64	-.44
5 005	29 40	1.29	.40	1.32	1.30	1.54	.80
6 006	27 40	.97	.39	1.17	1.13	.93	.46
7 007	27 40	.97	.39	1.18	1.13	.97	.47
8 008	29 40	1.29	.40	1.11	1.00	.61	.13
9 009	28 40	1.13	.40	1.55	1.72	2.53	1.65
10 010	29 40	1.29	.40	1.53	1.82	2.39	1.71
11 011	30 40	1.45	.41	1.21	1.06	1.04	.29
12 012	27 40	.97	.39	1.36	1.37	1.80	1.03
13 013	28 40	1.13	.40	1.56	1.71	2.57	1.62
14 014	32 40	1.82	.44	1.27	1.26	1.15	.62
15 015	32 40	1.82	.44	1.44	1.78	1.75	1.34
16 016	26 40	.82	.38	1.21	1.12	1.14	.45
17 017	32 40	1.82	.44	1.22	.97	.98	.12
18 018	32 40	1.82	.44	1.07	.82	.38	-.16
19 019	35 40	2.48	.51	1.04	.77	.23	-.05
20 020	35 40	2.48	.51	1.16	.86	.60	.07
21 021	27 40	.97	.39	1.14	1.53	.80	1.37
22 022	32 40	1.82	.44	1.27	1.35	1.17	.76
23 023	26 40	.82	.38	1.13	1.10	.76	.39
24 024	28 40	1.13	.40	1.01	1.00	.12	.12
25 025	27 40	.97	.39	1.16	1.17	.91	.57
26 026	21 40	.12	.37	1.54	1.47	2.76	1.55
27 027	23 39	.45	.38	1.55	1.75	2.69	2.11

28 028		26 40		.82	.38		1.22	1.53	1.18	1.45
29 029		29 40		1.29	.40		1.17	1.21	.90	.60
30 030		29 40		1.29	.40		1.23	1.29	1.15	.76
31 031		28 40		1.13	.40		1.04	.89	.25	-.17
32 032		27 40		.97	.39		1.01	.86	.10	-.29
33 033		27 40		.97	.39		1.01	.86	.10	-.29
34 034		29 40		1.29	.40		1.47	1.62	2.15	1.38
35 035		29 40		1.29	.40		.74	.72	-1.41	-.60
36 036		27 40		.97	.39		.77	.63	-1.26	-1.06
37 037		27 40		.97	.39		.80	.65	-1.11	-.98
38 038		27 40		.97	.39		1.32	2.38	1.65	2.88
39 039		36 40		2.77	.56		1.40	2.73	1.09	1.62
40 040		14 40		-.84	.38		1.45	1.93	2.38	2.22
41 041		26 40		.82	.38		.73	.70	-1.61	-.87
42 042		14 29		-.07	.44		1.12	1.00	.64	.11
43 043		21 37		.26	.39		1.16	1.09	.87	.38
44 044		24 40		.54	.38		.54	.44	-3.07	-2.22
45 045		25 40		.68	.38		.66	.56	-2.11	-1.51
46 046		28 40		1.13	.40		.71	.53	-1.62	-1.34
47 047		28 40		1.13	.40		.60	.45	-2.39	-1.67
48 048		29 40		1.29	.40		1.39	1.72	1.83	1.55
49 049		27 40		.97	.39		.62	.57	-2.31	-1.29
50 050		26 40		.82	.38		.69	.57	-1.87	-1.41
51 051		27 37		1.23	.43		.57	.41	-2.26	-1.60
52 052		25 40		.68	.38		.67	.58	-2.06	-1.42
53 053		28 40		1.13	.40		.62	.46	-2.26	-1.62
54 054		27 40		.97	.39		.63	.58	-2.21	-1.25
55 055		28 40		1.13	.40		.91	.77	-.41	-.51
56 056		28 40		1.13	.40		.67	.50	-1.94	-1.46

*****Output Continues*****

Case Estimates In input Order

7/ 5/15 5:43

all on all (N = 167 L = 40 Probability Level= .50)

NAME	SCORE	MAXSCR	ESTIMATE	ERROR	INFIT	OUTFT	INFT	OUTFT
			MNSQ	MNSQ	t	t		
57 057		31 40		1.63 .43		.83 .56	-.74	-.88
58 058		28 40		1.13 .40		1.02 .92	.15	-.09
59 059		27 40		.97 .39		.66 .54	-2.06	-1.43
60 060		29 40		1.29 .40		1.26 1.29	1.27	.78
61 061		27 40		.97 .39		.84 .76	-.88	-.58
62 062		13 40		-.99 .38		1.77 2.36	3.68	2.79
63 063		14 40		-.84 .38		1.61 2.58	3.07	3.31
64 064		14 40		-.84 .38		1.52 1.85	2.66	2.07
65 065		26 40		.82 .38		.72 .70	-1.65	-.86
66 066		26 40		.82 .38		.53 .42	-3.12	-2.11
67 067		18 40		-.28 .37		1.36 1.32	1.98	1.09
68 068		13 38		-.96 .39		1.25 1.58	1.43	1.50
69 069		25 40		.68 .38		.57 .45	-2.82	-2.05
70 070		20 40		-.01 .37		.64 .54	-2.37	-1.87
71 071		15 40		-.70 .38		1.61 2.18	3.10	2.82
72 072		24 40		.54 .38		1.20 1.28	1.11	.93
73 073		9 40		-1.63 .42		1.11 2.34	.56	2.09
74 074		22 40		.26 .37		1.29 1.39	1.59	1.31
75 075		24 40		.54 .38		.52 .43	-3.29	-2.32
76 076		19 40		-.15 .37		.88 .82	-.70	-.57
77 077		22 40		.26 .37		.76 .74	-1.43	-.89
78 078		20 39		.01 .37		1.52 1.91	2.68	2.64
79 079		24 40		.54 .38		.98 1.18	-.08	.66
80 080		27 40		.97 .39		.91 1.08	-.45	.32
81 081		27 40		.97 .39		.77 .79	-1.28	-.50
82 082		20 40		-.01 .37		1.23 1.25	1.29	.91
83 083		23 40		.40 .37		.88 .90	-.65	-.23
84 084		21 40		.12 .37		1.33 1.41	1.79	1.37

85 085		27 40		.97	.39		1.38	1.42	1.90	1.14
86 086		25 40		.68	.38		.84	.68	-.90	-1.02
87 087		25 40		.68	.38		1.38	1.56	1.96	1.59
88 088		28 40		1.13	.40		.76	.57	-1.33	-1.16
89 089		30 40		1.45	.41		1.12	1.55	.63	1.19
90 090		26 40		.82	.38		.95	.93	-.22	-.10
91 091		27 40		.97	.39		1.03	.98	.22	.06
92 092		25 40		.68	.38		1.35	1.77	1.84	2.06
93 093		22 40		.26	.37		1.57	2.05	2.89	2.90
94 094		25 40		.68	.38		1.25	1.53	1.35	1.53
95 095		35 40		2.48	.51		.97	.61	.02	-.32
96 096		25 40		.68	.38		1.42	1.47	2.16	1.38
97 097		25 40		.68	.38		1.67	2.14	3.19	2.80
98 098		33 40		2.02	.46		1.17	1.16	.72	.45
99 099		31 40		1.63	.43		1.43	2.69	1.81	2.50
100 100		30 40		1.45	.41		1.13	1.20	.69	.56
101 101		31 40		1.63	.43		.86	.57	-.59	-.86
102 102		22 40		.26	.37		1.34	1.42	1.81	1.40
103 103		25 40		.68	.38		.75	.72	-1.45	-.85
104 104		20 40		-.01	.37		1.34	1.26	1.86	.94
105 105		22 40		.26	.37		.96	.86	-.17	-.42
106 106		24 39		.63	.38		1.05	.92	.32	-.15
107 107		24 40		.54	.38		.81	.67	-1.10	-1.11
108 108		29 39		1.38	.42		1.06	.89	.35	-.10
109 109		28 40		1.13	.40		.75	.60	-1.36	-1.06
110 110		24 40		.54	.38		.74	.63	-1.60	-1.29
111 111		27 40		.97	.39		.70	.53	-1.73	-1.43
112 112		25 40		.68	.38		.77	.61	-1.35	-1.31

*****Output Continues*****

Case Estimates In input Order

7/ 5/15 5:43

all on all (N = 167 L = 40 Probability Level= .50)

NAME	SCORE MAXSCR	ESTIMATE	ERROR	INFIT	OUTFT	INFT	OUTFT
			MNSQ	MNSQ	t	t	
113 113		22 40		.26	.37		.86 .76 -.81 -.82
114 114		23 40		.40	.37		.82 .73 -1.05 -.91
115 115		26 40		.82	.38		.87 .78 -.68 -.57
116 116		22 40		.26	.37		.74 .66 -1.60 -1.25
117 117		24 40		.54	.38		.70 .59 -1.86 -1.48
118 118		15 38		-.67	.38		1.00 1.16 .07 .57
119 119		25 40		.68	.38		.98 .81 -.04 -.51
120 120		22 40		.26	.37		.68 .56 -2.03 -1.71
121 121		25 40		.68	.38		.88 .72 -.67 -.87
122 122		24 40		.54	.38		1.13 1.15 .79 .56
123 123		26 40		.82	.38		.76 .62 -1.41 -1.20
124 124		26 40		.82	.38		.73 .59 -1.62 -1.30
125 125		25 40		.68	.38		1.13 1.11 .75 .43
126 126		22 39		.28	.38		.61 .50 -2.61 -2.00
127 127		24 40		.54	.38		.87 .73 -.71 -.89
128 128		26 40		.82	.38		1.01 .89 .14 -.21
129 129		29 40		1.29	.40		.92 .74 -.36 -.55
130 130		28 40		1.13	.40		.78 .63 -1.17 -.95
131 131		36 40		2.77	.56		1.00 .81 .11 .08
132 132		28 40		1.13	.40		.99 1.02 .01 .19
133 133		26 40		.82	.38		.74 .63 -1.52 -1.15
134 134		22 40		.26	.37		.83 .72 -.99 -.97
135 135		24 40		.54	.38		.80 .68 -1.14 -1.07
136 136		30 40		1.45	.41		.69 .50 -1.63 -1.20
137 137		32 40		1.82	.44		.74 .50 -1.16 -.94
138 138		27 40		.97	.39		1.07 1.10 .42 .38
139 139		29 40		1.29	.40		1.17 1.31 .91 .80
140 140		25 40		.68	.38		.61 .48 -2.53 -1.91

141	141		28	40		1.13	.40		.72	.54	-1.59	-1.31
142	142		30	40		1.45	.41		.81	.58	-.90	-.92
143	143		24	40		.54	.38		.75	.73	-1.51	-.89
144	144		16	40		-.56	.37		.79	.67	-1.27	-1.11
145	145		21	40		.12	.37		.62	.53	-2.49	-1.94
146	146		29	40		1.29	.40		.91	.69	-.41	-.69
147	147		25	40		.68	.38		.75	.61	-1.46	-1.31
148	148		25	40		.68	.38		.79	.68	-1.19	-1.02
149	149		28	40		1.13	.40		.95	.73	-.20	-.63
150	150		24	40		.54	.38		.68	.65	-1.99	-1.20
151	151		21	40		.12	.37		.78	.74	-1.33	-.93
152	152		23	39		.42	.38		.83	.77	-.97	-.73
153	153		27	40		.97	.39		.76	.64	-1.37	-1.01
154	154		28	40		1.13	.40		.73	1.17	-1.54	.55
155	155		24	40		.54	.38		1.06	1.03	.38	.20
156	156		36	40		2.77	.56		.96	.71	.01	-.05
157	157		27	40		.97	.39		.71	.60	-1.71	-1.17
158	158		19	40		-.15	.37		.90	.80	-.56	-.67
159	159		22	40		.26	.37		.88	.82	-.65	-.58
160	160		28	40		1.13	.40		.80	.61	-1.06	-1.03
161	161		29	40		1.29	.40		1.00	.74	.05	-.54
162	162		24	40		.54	.38		.93	.87	-.35	-.35
163	163		20	40		-.01	.37		.79	.68	-1.26	-1.17
164	164		25	40		.68	.38		.82	.69	-1.02	-.97
165	165		27	40		.97	.39		.72	.59	-1.63	-1.21
166	166		27	40		.97	.39		.67	.51	-1.95	-1.53
167	167		26	40		.82	.38		.74	.61	-1.52	-1.23

*****Output Continues*****

Case Estimates In input Order

7/ 5/15 5:43

all on all (N = 167 L = 40 Probability Level= .50)

NAME		SCORE MAXSCR		ESTIMATE	ERROR		INFIT	OUTFT	INFT	OUTFT
				MNSQ	MNSQ	t	t			

Mean				.82		1.00	1.01	-.09	-.06	
SD				.69		.29	.50	1.57	1.27	

Tabel 3. Hasil Analisis Quest Soal Paket C

Hasil penelitian menunjukkan bahwa persamaan linear equating dari 4 kelompok hasil UAS Fisika 2015 (Paket A, Paket B, Paket C, dan Paket D), yang digunakan, yaitu;

1. $Y' = a(X-c) + d$, maka hasil yang diperoleh untuk equating dari paket A (X) ke Paket B (Y) adalah;

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

$$a = \frac{5,345}{4,235} = 1,262 \quad \text{dan } c = 20,75 \quad \text{serta } d = 23,25$$

$$Y' = 1,262 (X - 20,75) + 23,25$$

$$Y' = 1,262 X - 2.9365 \text{ untuk equating dari Paket A ke B}$$

2. $Y' = a(X-c) + d$, maka hasil yang diperoleh untuk equating dari paket B (X) ke Paket C (Y) adalah;

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

$$a = \frac{4,236}{3,965} = 1,068 \quad \text{dan } c = 21,15 \quad \text{serta } d = 19,45$$

$$Y' = 1,068 (X - 21,15) + 19,45$$

$$Y' = 1,068 X - 15,122 \text{ untuk equating dari Paket B ke C}$$

3. $Y' = a(X-c) + d$, maka hasil yang diperoleh untuk equating dari paket C (X) ke Paket D (Y) adalah;

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

$$a = \frac{5,205}{4,145} = 1,256 \quad \text{dan } c = 18,65 \quad \text{serta } d = 22,25$$

$$Y' = 1,256 (X - 18,65) + 22,25$$

$$Y' = 1,256 X - 1,174 \text{ untuk equating dari Paket C ke A}$$

4. $Y' = a(X-c) + d$, maka hasil yang diperoleh untuk equating dari paket D (X) ke Paket A (Y) adalah;

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

$$a = \frac{4,325}{3,125} = 1,384 \quad \text{dan } c = 12,34 \quad \text{serta } d = 16,23$$

$$Y' = 1,384 (X - 12,34) + 16,23$$

$$Y' = Y' = 1,384 X - 0,932 \text{ untuk equating dari Paket D ke A}$$

5. $Y' = a(X-c) + d$, maka hasil yang diperoleh untuk equating dari paket A (X) ke Paket C (Y) adalah;

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

$$a = \frac{6,243}{5,145} = 1,214 \quad \text{dan } c = 16,32 \quad \text{serta } d = 20,42$$

$$Y' = 1,214 (X - 16,32) + 20,42$$

$$Y' = 1,214 X - 0,607 \text{ untuk equating dari Paket A ke C}$$

6. $Y' = a(X-c) + d$, maka hasil yang diperoleh untuk equating dari paket B (X) ke Paket C (Y) adalah;

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

$$a = \frac{4,765}{3,876} = 1,229 \quad \text{dan } c = 14,44 \quad \text{serta } d = 12,76$$

$$Y' = 1,229 (X - 14,44) + 12,76$$

$$Y' = 1,229 X - 4,987 \text{ untuk equating dari Paket B ke C}$$

7. $Y' = a(X-c) + d$, maka hasil yang diperoleh untuk equating dari paket B (X) ke Paket D (Y) adalah;

$$a = \frac{\sigma_x}{\sigma_y} \quad \text{dan } c = \mu_x \quad \text{serta } d = \mu_y$$

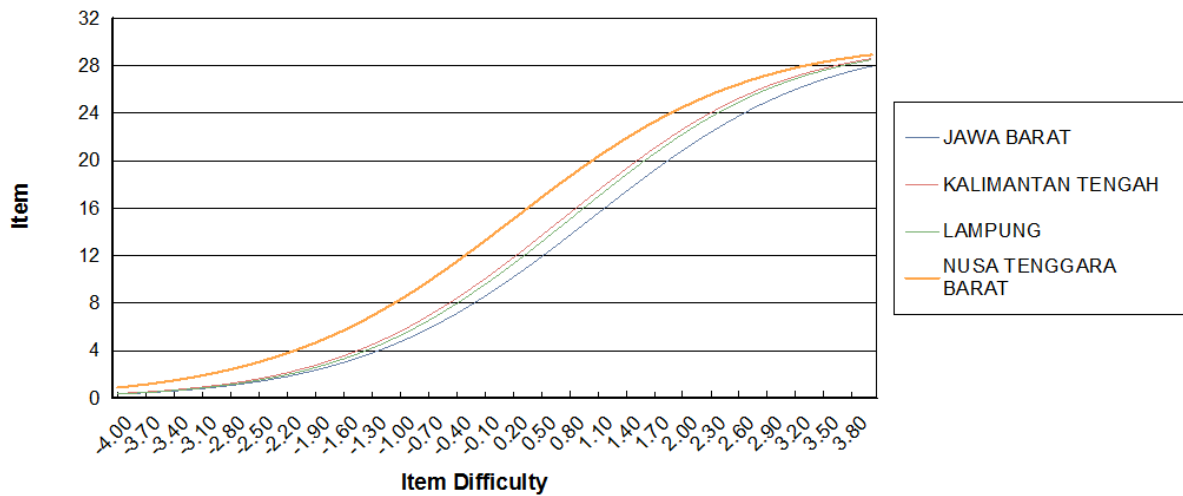
$$a = \frac{6,865}{4,987} = 1,256 \quad \text{dan } c = 16,72 \quad \text{serta } d = 19,34$$

$$Y' = 1,256 (X - 16,72) + 19,34$$

$$Y' = 1,256 X - 1,660. \text{ untuk equating dari Paket B ke D}$$

Hasil rekapitulasi karakteristik butir Model rasch 1 Parameter (tingkat kesukaran butir soal) dari keempat paket soal yang berasal dari soal non Anchor (30 soal yang dibuat oleh MGMP dan atau guru),

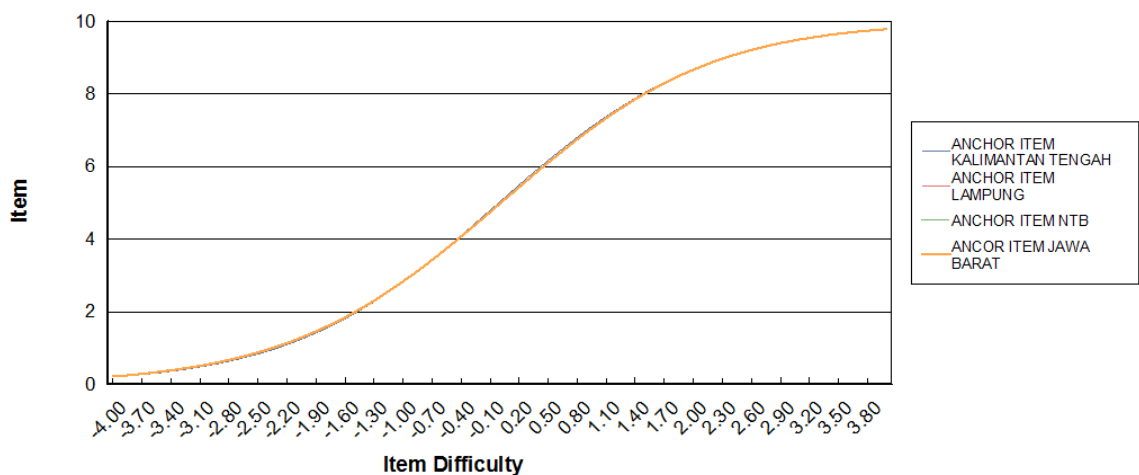
dapat dilihat pada grafik di bawah ini.



Gambar 4.1. Hasil Equating Soal Non Anchor Untuk 4 Paket Soal Berbeda (30 butir soal)

Dari Gambar terlihat untuk soal yang dibuat oleh MGMP dan atau guru, terlihat urutan tingkat kesukaran soal adalah Jawa Barat, Lampung, Kalimantan Tengah dan Nusa Tenggara Barat. Artinya soal di wilayah Jawa Barat relatif lebih sulit dibandingkan dengan ketiga wilayah lain.

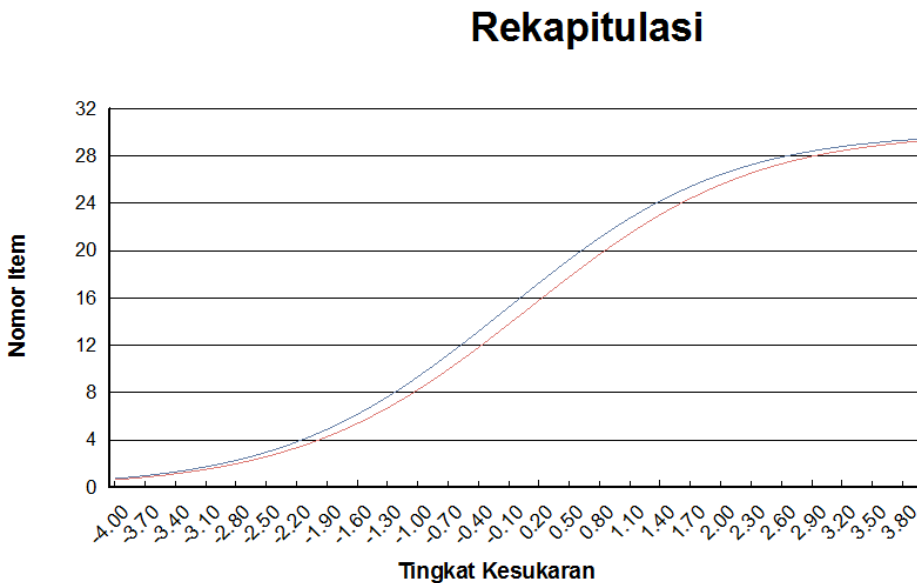
Hasil rekapitulasi karakteristik butir Model rasch 1 Parameter (tingkat kesukaran butir soal) dari keempat paket soal yang berasal dari soal Anchor (10 soal yang dibuat oleh tim peneliti), dapat dilihat pada grafik di bawah ini.



Gambar 4.2. Hasil Equating Soal Anchor Untuk 4 Paket Soal Sama (10 butir soal)

Dari Gambar terlihat untuk soal yang dibuat oleh tim peneliti, memiliki karakteristik butir yang relatif sama, khususnya bila ditinjau dari parameter tingkat kesukaran butir soal. Artinya karakteristik butir soal anchor tidak terpengaruh oleh karakteristik responden, dan hal ini sesuai dengan teori respon butir soal.

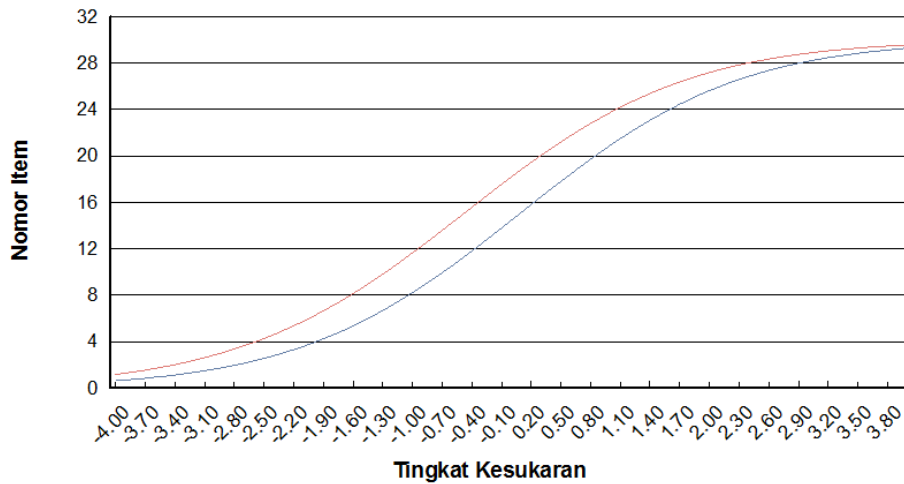
Adapun rekapitulasi hasil equating Paket A (X) kepada Paket B (Y) maka diperoleh bentuk grafik / kurva karakteristik equating dari hasil perhitungan di atas adalah sebagai berikut:



Gambar 4.3. Grafik rekapitulasi hasil equating Paket A dan B

Adapun rekapitulasi hasil equating Paket B (X) kepada Paket C (Y) maka diperoleh bentuk grafik / kurva karakteristik equating dari hasil perhitungan di atas adalah sebagai berikut:

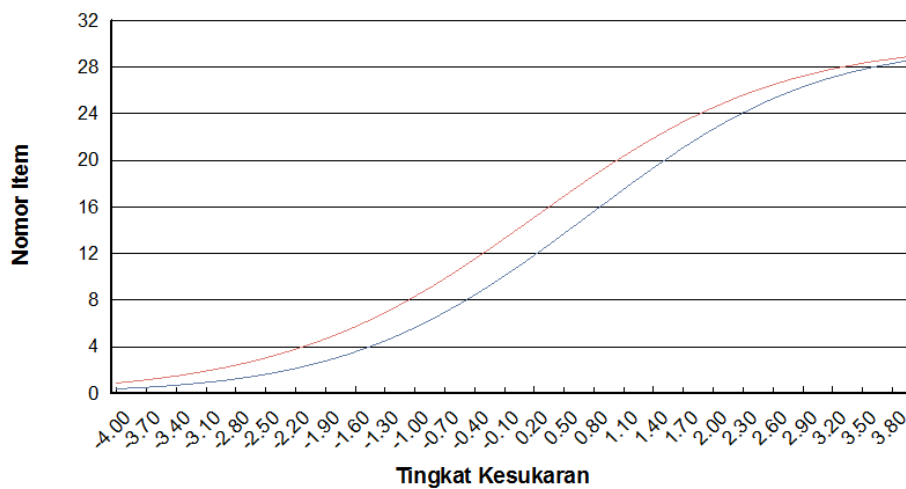
Rekapitulasi



Gambar 4.4. Grafik rekapitulasi hasil equating Paket B dan C

Adapun rekapitulasi hasil equating Paket C (X) kepada Paket A (Y) maka diperoleh hasil bentuk grafik / kurva karakteristik equating dari hasil perhitungan di atas adalah sebagai berikut:

Rekapitulasi



Gambar 4.5. Grafik rekapitulasi hasil equating Paket A dan C

Dari gambar di atas, dipelihatkan bahwa garis hasil linear equating paket A terhadap paket patokan B sama nilai-nilai rata-ratanya, hal itulah memang yang menjadi dasar pada rumus linear equating. Namun hasil linear equating bagi skor yang rendah berada dibawah nilai patokan, sementara skor yang lebih tinggi akan berada diatas nilai patokan hal itu disebabkan karena proses equating dilakukan dari paket yang sukar kepada paket yang mudah. Bila proses equating tersebut dari paket mudah ke paket yang sukar maka garis linear hasil penyetaraan akan sebaliknya

Berdasarkan persamaan pertama penyetaraan linier tersebut tergambar bahwa kedua paket soal Ujian Sekolah Fisika di SMAN (Paket A) dan SMAN di Lampung (Paket B) memiliki tingkat kesukaran yang berbeda. Paket B lebih sukar dibandingkan paket A. Maka, persamaan linear equating yang terbaik terhadap skor akhir yang tidak merugikan peserta ujian adalah penyetaraan dari paket sukar ke paket mudah dengan persamaan $Y' = 1,262 X - 2.9365$.

Jika paket B dan paket A tidak setara dalam artian bahwa paket tes tersebut tidak mencerminkan alat ukur yang benar-benar reliabel, padahal ia diberikan di satu sekolah atau antar sekolah dalam satu kabupaten, maka hasil akhir dari ujian harian tersebut akan terjadi kesenjangan. Siswa dengan kemampuan tinggi yang mendapatkan paket sukar akan dirugikan dengan hasil tes tersebut, sedangkan yang mendapat paket mudah tentunya akan diuntungkan. Dengan demikian hasil dari ujian tersebut tidak mencerminkan prestasi siswa yang sebenarnya dalam satu kabupaten karena bebedanya paket-paket yang diujikan. Selain itu dalam menafsirkan hasil pengukuran, dari beberapa paket Ujian Sekolah sering diperlakukan sama, padahal tingkat kesukaran berbeda. Hal demikian dapat dikatakan bahwa pengukuran tersebut menyimpang dari teori pengukuran yang seharusnya reliabel untuk satu populasi yang sama.

Berdasarkan uraian di atas seharusnya kedua paket tersebut dilakukan penyetaraan hasil skornya, dan itulah yang perlu dilakukan hasil ujian pada perangkat tes soal Ujian Sekolah Fisika antara paket B dan paket A. Bila paket yang diberikan berbeda padahal mengukur kontrak yang sama dan dari kisi-kisi yang sama, tidak adil jika hasil tes tidak disetarakan kecuali ada konversi sebelumnya yang telah dirancang dan disiapkan untuk menyesuaikan dengan perbedaan paket dan perbedaan tingkat kesukarannya.

Berdasarkan persamaan kedua penyetaraan linier tersebut tergambar bahwa kedua paket soal harian Fisika di Lampung (Paket B) dan SMAN Palangkaraya (Paket C) memiliki tingkat kesukaran

yang berbeda. Paket B lebih sukar dibandingkan paket C. Maka, persamaan linear equating yang terbaik terhadap skor akhir yang tidak merugikan peserta ujian adalah penyetaraan dari paket sukar ke paket mudah dengan persamaan $Y' = 1,068 X - 15,122$.

Jika paket B dan paket C tidak setara dalam artian bahwa paket tes tersebut tidak mencerminkan alat ukur yang benar-benar reliabel, padahal ia diberikan di satu sekolah atau antar sekolah dalam satu kabupaten, maka hasil akhir dari ujian harian tersebut akan terjadi kesenjangan. Siswa dengan kemampuan tinggi yang mendapatkan paket sukar akan dirugikan dengan hasil tes tersebut, sedangkan yang mendapat paket mudah tentunya akan diuntungkan. Dengan demikian hasil dari ujian tersebut tidak mencerminkan prestasi siswa yang sebenarnya dalam satu kabupaten karena bebedanya paket-paket yang diujikan. Selain itu dalam menafsirkan hasil pengukuran, dari beberapa paket Ujian Sekolah sering diperlakukan sama, padahal tingkat kesukaran berbeda. Hal demikian dapat dikatakan bahwa pengukuran tersebut menyimpang dari teori pengukuran yang seharusnya reliabel untuk satu populasi yang sama.

Berdasarkan uraian di atas seharusnya kedua paket tersebut dilakukan penyetaraan hasil skornya, dan itulah yang perlu dilakukan hasil ujian pada perangkat tes soal Ujian Sekolah Fisika antara paket C dan paket A. Bila paket yang diberikan berbeda padahal mengukur konstruk yang sama dan dari kisi-kisi yang sama, tidak adil jika hasil tes tidak disetarakan kecuali ada konversi sebelumnya yang telah dirancang dan disiapkan untuk menyesuaikan dengan perbedaan paket dan perbedaan tingkat kesukarannya.

Berdasarkan persamaan ketiga penyetaraan linier tersebut tergambar bahwa kedua paket soal UAS Fisika di SMAN Lombok Timur (Paket A) dan SMAN Palangkaraya (Paket C) memiliki tingkat kesukaran yang berbeda. Paket A lebih sukar dibandingkan paket C. Maka, persamaan linear equating yang terbaik terhadap skor akhir yang tidak merugikan peserta ujian adalah penyetaraan dari paket sukar ke paket mudah dengan persamaan $Y' = 1,256 X - 1,174$.

Jika paket C dan paket A tidak setara dalam artian bahwa paket tes tersebut tidak mencerminkan alat ukur yang benar-benar reliabel, padahal ia diberikan di satu sekolah atau antar sekolah dalam satu kabupaten, maka hasil akhir dari ujian harian tersebut akan terjadi kesenjangan. Siswa dengan kemampuan tinggi yang mendapatkan paket sukar akan dirugikan dengan hasil tes tersebut, sedangkan yang mendapat paket mudah tentunya akan diuntungkan. Dengan demikian hasil dari ujian tersebut tidak mencerminkan prestasi siswa yang sebenarnya dalam satu kabupaten karena bebedanya paket-

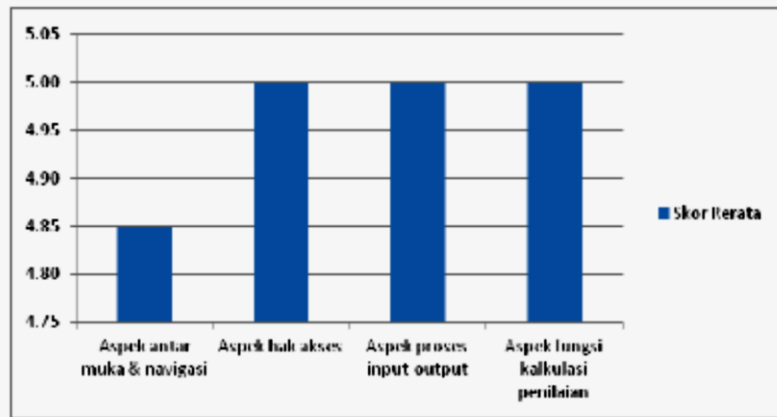
paket yang diujikan. Selain itu dalam menafsirkan hasil pengukuran, dari beberapa paket Ujian Sekolah sering diperlakukan sama, padahal tingkat kesukaran berbeda. Hal demikian dapat dikatakan bahwa pengukuran tersebut menyimpang dari teori pengukuran yang seharusnya reliabel untuk satu populasi yang sama.

Berdasarkan uraian di atas seharusnya kedua paket tersebut dilakukan penyetaraan hasil skornya, dan itulah yang perlu dilakukan hasil ujian pada perangkat tes soal Ujian Sekolah Fisika antara paket C dan paket A. Bila paket yang diberikan berbeda padahal mengukur kontrak yang sama dan dari kisi-kisi yang sama, tidak adil jika hasil tes tidak disetarakan kecuali ada konversi sebelumnya yang telah dirancang dan disiapkan untuk menyesuaikan dengan perbedaan paket dan perbedaan tingkat kesukarannya.

C. Pengujian Kualitatif

Pengujian kualitatif dilakukan dengan menggunakan metode angket teknik *simple random sampling onliner* (sampel *onliner* acak sederhana). Kuesioner online yang telah dibuat dibagikan kepada civitas akademik yang biasa menggunakan internet dan aplikasi dokumen. Terdapat 5 variabel penilaian pada kuesioner ini yaitu variabel kemudahan akses, kemudahan penggunaan, daya tarik, fitur-fitur yang diberikan, dan tampilan (*interface*) aplikasi. Setiap variabel diberikan skala penilaian 1-5 dengan skala 1 untuk respon “Sangat Buruk” dan skala 5 untuk respon “Sangat Baik”.

Berdasarkan hasil kuesioner didapatkan bahwa aplikasi dokumen untuk variabel kemudahan akses, kemudahan penggunaan, daya tarik aplikasi, fitur-fitur yang diberikan, dan tampilan (*interface*) aplikasi rata-rata telah mendekati skala respon sangat baik. Ini berarti aplikasi dokumen berbasis virtual ini telah dinyatakan layak digunakan untuk menunjang aktivitas dokumen civitas akademik. Namun terdapat pula beberapa masukan dari para responden yang membangun seperti menyarankan agar tampilan dibuat lebih menarik, penambahan tutorial penggunaan, penambahan fitur lainnya serta penggunaan domain dan IP Public untuk mempercepat akses aplikasi. Grafik Hasil kuesioner variabel kemudahan akses aplikasi dapat dilihat pada Gambar 9.



Gambar 8. Kuisisioner tentang Kemudahan Akses oleh User (guru)

D. Hasil Analisis Ketidakwaajaran Skor

Indeks ketidakwaajaran sekor berdasarkan teori respon butir yang digunakan berdasarkan hasil analisis program Quest diantaranya adalah ketidak wajaran melalui kebolehjadian dan indeks ketidakwaajaran residu bakuan terkuadrat (Naga, 1992: 164-168).

1. Kewajaran Melalui Kebolehjadian

Indeks kewajaran sekor ditentukan melalui teori responsi butir. Karena estimasi parameter dilakukan melalui kebolehjadian maksimum, maka indeks kewajaran dihitung melalui kebolehjadian. Tingginya nilai kebolehjadian dijadikan indeks kewajaran; makin tinggi kebolehjadian makin wajar sekor responden. Di dalam proses perhitungan digunakan logaritma, mencakup; indeks kewajaran l_0 , indeks kewajaran l_g , dan indeks kewajaran l_z .

(a) Indeks Kewajaran Kebolehjadian l_0

Indeks kewajaran ini menggunakan logaritma dari kebolehjadian. Kebolehjadian pada θ yang diestimasi melalui kebolehjadian maksimum

$$L(X | \theta) = \prod_{i=1}^N P_i(\theta)^{X_i} Q_i(\theta)^{1-X_i} \quad (1)$$

dengan jawaban betul $X_i = 1$ dan jawaban salah $X_i = 0$.

Indeks kewajaran l_0

$$\begin{aligned}
l_0 &= \ln L(X | \theta) \\
&= \sum_{i=1}^N [X_i \ln P_i(\theta) + (1 - X_i) \ln Q_i(\theta)]
\end{aligned} \tag{2}$$

dengan nilai $l_0 \leq 0$. Karena telah digunakan θ yang diperoleh melalui kebolehjadian maksimum, maka pada sekor wajar seharusnya makin tinggi l_0 makin baik. Nilai l_0 yang rendah sekali menunjukkan ketidakwajaran sekor. Jika butir mudah dijawab betul dan butir sukar dijawab salah, maka indeks kewajaran akan tinggi. Jika butir mudah dijawab salah dan butir sukar dijawab betul, maka indeks kewajaran akan rendah.

(b) Indeks Kewajaran Kebolehjadian l_g

Indeks kewajaran l_g mereratakan indeks kewajaran berdasarkan butir yang dijawab, sehingga menjadi

$$l_g = e^{\frac{l_0}{N}} \tag{3}$$

dengan N = banyaknya butir yang dijawab. Karena perhitungan didasarkan pada indeks per butir yang dijawab, maka terdapat perlakuan sama di antara responden yang menjawab banyak butir dan yang sedikit butir. Makin tinggi nilai indeks kewajaran makin wajar sekor responden

(c). Indeks Kewajaran Kebolehjadian Nilai Baku l_z

Apabila kemampuan responden θ berbeda, maka indeks kewajaran l_g menjadi kurang memadai. Untuk mengatasi hal ini, digunakan indeks kewajaran nilai baku.

$$l_z = \frac{l_0 - \mu_{l_0}}{\sigma_{l_0}} \tag{4}$$

Perhitungan indeks kewajaran memerlukan nilai rerata dan simpangan baku pada l_0 .

$$\mu_{l_0} = \frac{\sum_{i=1}^N l_{0i}}{N} = \sum_{i=1}^N [P_i(\theta) \ln P_i(\theta) + Q_i(\theta) \ln Q_i(\theta)] \tag{5}$$

Melalui substitusi $m_i(\theta) = P_i(\theta) \ln P_i(\theta) + Q_i(\theta) \ln Q_i(\theta)$ maka rerata menjadi

$$\mu_{l_0} = \sum_{i=1}^N m_i(\theta) \tag{6}$$

Simpangan bakunya adalah:

$$\begin{aligned}\sigma_{l_0} &= \sqrt{\frac{N \sum_{i=1}^N l_{0_i}^2 - \left(\sum_{i=1}^N l_{0_i} \right)^2}{N^2}} \\ &= \sqrt{\sum_{i=1}^N \left[P_i(\theta) Q_i(\theta) \left(\ln \frac{P_i(\theta)}{Q_i(\theta)} \right)^2 \right]}\end{aligned}\quad (7)$$

2. Indeks Kewajaran Residu Bakuan Terkuadrat

Responden menghasilkan jawaban berupa jawaban betul $X_i = 1$ dan jawaban salah $X_i = 0$. Misalnya untuk model logistik menghasilkan probabilitas betul $P_i(\theta)$ dan probabilitas salah $Q_i(\theta)$.

Selisih di antara mereka adalah residu R_i , yaitu:

$$R_i = X_i - P_i(\theta) \quad (8)$$

Residu menjadi dasar untuk menunjukkan kewajaran sekor responden rerata dan simpangan baku sebagai berikut:

$$\text{Rerata} \quad \mu_{X_i} = P_i(\theta) \quad (9)$$

$$\text{Simpangan baku} \quad \sigma_{X_i} = \sqrt{P_i(\theta) Q_i(\theta)} \quad (10)$$

Nilai baku selisih atau residunya adalah:

$$S_{R_i} = \frac{X_i - \mu_{X_i}}{\sigma_{X_i}} = \frac{X_i - P_i(\theta)}{\sqrt{P_i(\theta) Q_i(\theta)}} \quad (11)$$

Pada saat $X_i = 0$

$$S_{R_i}(X_i = 0) = -\sqrt{\frac{P_i(\theta)}{Q_i(\theta)}} \quad (12)$$

Pada saat $X_i = 1$

$$S_{R_i}(X_i = 1) = \sqrt{\frac{Q_i(\theta)}{P_i(\theta)}} \quad (13)$$

Indeks kewajaran sekor terkuadrat untuk N butir

$$W = \sum_{i=1}^N S_{R_i}^2 = \sum_{i=1}^N \left[X_i \frac{Q_i(\theta)}{P_i(\theta)} - (1 - X_i) \frac{P_i(\theta)}{Q_i(\theta)} \right] \quad (14)$$

Pada model logistik L1P

$$\frac{Q_i(\theta)}{P_i(\theta)} = e^{-D(\theta-b_i)} \quad \frac{P_i(\theta)}{Q_i(\theta)} = e^{D(\theta-b_i)} \quad (15)$$

Sehingga

$$W = \sum_{i=1}^N \left[X_i e^{-D(\theta-b_i)} - (1 - X_i) e^{D(\theta-b_i)} \right] \quad (16)$$

W diturunkan dari residu sehingga makin besar W makin besar residu dan makin tidak wajar sekor responden.

Untuk menggambarkan bagaimana pola jawaban responden dihubungkan dengan kemampuan (*ability*) dan karakteristik butir soal seperti tingkat kesukaran digunakan *person-fit statistic*. *Person-fit statistic* dapat dijadikan sebagai indikasi kewajaran sekor peserta ujian. Wright dan Linacre (1992) mengungkapkan indeks ketidakwajaran dengan menggunakan *person-fit* dan menamakannya *weighted total fit mean square* dan *unweighted total fit mean square* (Rudner, 1983).

Person-fit statistic dapat diperoleh dengan melihat *outfit statistic* pada output Quest. *Outfit statistic person* menunjukkan bagaimana perilaku yang tidak diharapkan dari soal yang mempunyai tingkat kesukaran jauh dengan kemampuan peserta ujian tersebut (Setiadi, 1999). Dengan kata lain, peserta ujian dengan kemampuan tinggi tidak dapat menjawab soal dengan tingkat kesukaran rendah. Atau sebaliknya, peserta dengan kemampuan yang rendah tetapi menjawab benar butir soal dengan tingkat kesukaran tinggi. Hal ini sejalan dengan pengertian ketidakwajaran.

Untuk mempertajam analisis terkait indeks ketidakwajaran digunakan metode SHL diambil dari nama Sato, Harnisch dan Linn, yang selanjutnya disebut indeks kehati-hatian (*caution index*). Metode ini menggunakan banyaknya jawaban salah pada butir mudah dan banyaknya jawaban benar pada butir sulit. Selanjutnya dicari index kehati-hatian dalam bentuk proporsi terhadap jawaban benar dari seluruh peserta. Dalam menentukan indeks tersebut, dibuat matriks jawaban peserta ujian. Karena fokus pada peserta ujian, maka pada bagian kolom pertama adalah peserta ujian dan pada bagian baris adalah butir soal. Selanjutnya, peserta diurutkan dari tinggi sampai rendah dimana peserta dengan tertinggi ditempatkan di atas. Butir soal juga diurutkan dari butir termudah sampai pada butir tertinggi. Langkah selanjutnya adalah

menentukan batas dari jawaban benar dan jawaban salah. Selanjutnya, yang perlu diperhatikan adalah pola jawaban peserta ujian. Berapa butir soal yang benar pada tempat yang seharusnya salah karena butir tersebut sukar. Juga berapa butir soal yang salah pada tempat yang seharusnya peserta tersebut benar karena butir tersebut mudah.

Beberapa notasi untuk rumus indeks kehati-hatian, adalah:

t = batas di antara jawaban salah dan jawaban betul jika responden berhati-hati

f_{gi} = sekor butir pada indeks kehati-hatian untuk responden ke- g

f_t = banyaknya butir di bawah batas t

N = banyaknya butir

X_{gi} = sekor butir oleh responden ke- g

= 1 untuk jawaban betul

= 0 untuk jawaban salah

Indeks kehati-hatian SHL untuk responden ke- g , dinyatakan dengan persamaan:

$$c_g = \frac{A_g - B_g}{C - D} \quad (17)$$

Dimana : $A_g = \text{sekor jawaban salah}$

$$= \sum_{i=1}^{f_t} (1 - X_{gi}) f_{gi} \quad (18)$$

$B_g = \text{sekor jawaban betul}$

$$= \sum_{i=f_t+1}^N X_{gi} f_{gi} \quad (19)$$

$$C = \sum_{i=1}^{f_t} f_{gi} \quad \text{dan} \quad D = \sum_{i=N-f_t+1}^N f_{gi} \quad (20)$$

Data yang diperoleh dianalisis secara matematis dengan bantuan komputer menggunakan program *Quest*. Analisis program *Quest* yang dibahas dalam penelitian ini adalah bagian *output* menurut teori respon butir, yang terdiri dari; kecocokan dengan model, tingkat kesukaran butir tes, reliabilitas soal, dan estimasi kemampuan responden. Estimasi butir dan responden dilakukan dengan prosedur PROX (*normal approximation estimation*). Kecocokan antara kemampuan responden (θ) dan indeks kesukaran butir (b) akan menghasilkan akurasi dalam pengukuran. Akurasi maksimal terjadi saat $P(\theta) = 0,5$. Estimasi

parameter dilakukan dengan membuang responden yang benar dan salah semua. Estimasi parameter responden dan butir dilakukan serentak karena keduanya belum diketahui. Estimasi terus dilakukan sampai nilai parameter responden dan butir konstan.

Parameter pertama adalah kecocokan butir dengan model Rasch, yaitu dengan melihat nilai *infit meansquare*. Penetapan *fit* testi (*case/person*) secara keseluruhan dengan program Quest (Adam & Kho, 1996) didasarkan pada besarnya nilai rata-rata INFIT *Mean of Square* (INFIT MNSQ) beserta simpangan bakunya. Penetapan *o* didasarkan pada besarnya nilai rata-rata INFIT *Mean of* INFIT *t*. Penetapan *fit* tiap testi (*case/person*) dengan model dalam program Quest didasarkan pada besarnya nilai INFIT MNSQ atau nilai INFIT *t* item yang bersangkutan. Langkah perhitungannya mengikuti langkah yang ditulis Wright & Masters (1982: 108-109). Tahap kedua adalah melihat nilai *outfit t*. Tahap ketiga adalah menganalisis indeks kesukaran butir dengan melihat *thresholds*. Butir soal yang cocok dengan model Rasch memiliki *infit meansquare* 0,77 -1,30 nilai *outfit t* ≤ 2 . Hampir semua soal dinyatakan cocok dengan model Rasch karena nilai *infit meansquare*-nya 0,77 -1,30. Soal nomor 14 dinyatakan gugur dengan nilai *outfit t*-nya 2,15

Parameter kedua yang dianalisis adalah tingkat kesukaran butir tes. Pada distribusi tingkat kesukaran soal dan kemampuan responden dapat dilihat pada file map. Bentangan tingkat kesukaran dan tingkat kemampuan siswa tersebut berada pada satu garis sehingga akan dapat diketahui posisi setiap subjek terhadap tingkat kesulitan butir yang dikerjakan. Tingkat kemampuan tes maupun tingkat kesukaran butir dalam *Rasch Model* (RM) diekspresikan pada satu garis berupa absis pada grafik dengan satuan berupa logit (*logg-odd unit*) (Keeves & Alagumalai, 1999: 27). Proporsi varians total dari estimasi skala untuk *person* sebesar β_n yang berasosiasi dengan varians parameter, ditentukan oleh besarnya indeks separasi *person* sebesar *S*. Indeks separasi *person* itulah yang dianggap sama dengan koefisien reliabilitas tes (Wright & Masters: 1982: 106). Namun, tetap harus diperhatikan bahwa perhitungan besarnya *error* pengukuran pada indeks sparasi *person* berbeda dengan perhitungan *error* varians pada CTT (Keeves & Masters, 1999: 275-276).

File ini menyajikan persebaran responden menurut tingkat kemampuannya dan persebaran butir menurut tingkat kesukarannya dalam logit -2,0 sampai +2,0. Dari 40 butir soal yang dibuat, terdapat 2 soal berkategori sangat mudah, 8 soal berkategori mudah, 25 soal berkategori sedang, 4 soal berkategori sukar dan 1 soal berkategori sangat sukar. Butir tes nomor 9 dianggap soal yang paling mudah dan nomor 16 sebagai soal yang paling sukar.

Reliabilitas soal sebagai parameter yang ketiga mempunyai nilai 0,97 sehingga tes dianggap mempunyai reliabilitas sangat tinggi. Realibilitas ini terlihat dari output program Quest yang menyajikan hasil realiabilitas tes menurut CTT, yakni berupa indeks konsistensi internal, yang untuk pesekor an politomus merupakan indeks alpha Cronbach dan untuk pensekor an dikotomus merupakan indeks KR-20 (Adam & Khoo, 1996: 93). Ini berarti bahwa dengan program QUEST dapat diperoleh item atau butir dan testi yang fit, disertai dengan reliabilitas instrumen tes tersebut

Langkah selanjutnya adalah melakukan estimasi kemampuan responden. Kemampuan responden dapat dilihat dari banyaknya butir yang dapat dijawab dengan benar. Semakin banyak butir yang dapat dijawab dengan benar, maka kemampuan responden semakin tinggi. Estimasi kemampuan responden dapat dilihat pada file *teta* pada nilai *estimate*. Sampel pada penelitian ini sebanyak 1363 responden peserta UAS mata pelajaran Fisika SMA r tahun 2015. Hasil analisis menunjukkan bahwa 254 responden (18,63 %) memiliki kemampuan tinggi, 753 responden (55,25 %) memiliki kemampuan sedang, dan 356 responden (26,12 %) memiliki kemampuan rendah. Peta butir dan Responden menggambarkan distribusi tingkat kesukaran butir dan kemampuan responden secara bersamaan dalam skala logit. Butir soal yang paling sukar yaitu butir item 23 yang berada di bagian teratas peta dengan tingkat kesukaran bernilai 2,27 sedangkan butir soal yang paling mudah adalah butir soal nomor 39 dengan tingkat kesukaran bernilai - 2,63. Kemampuan resonden tertinggi terlihat pada tanda silang teratas yaitu bernilai 2,23 dan terendah - 2,18. Dari analisis kualitas butir 40 soal pilihan ganda UAS mata pelajaran Fisika SMA tahun pelajaran 2014/2015 secara kuantitatif baik menurut pendekatan teori respon butir.

Setelah dilakukan analisis butir berdasarkan data empirik hasil UAS mata pelajaran Fisika SMA tahun 2015, langkah selanjutnya adalah menganalisis ketidakwajaran (*inappropriateness*) sekor tes tersebut. Ada dua teknik yang digunakan dalam menganalisis ketidakwajaran sekor pada penelitian ini, yaitu teknik itu adalah *personal biserial* dan *caution index*. Teknik personal biserial digunakan untuk melihat kesesuaian pola jawaban sekor individu pada semua butir soal dan dibandingkan dengan pola jawaban sekor kelompok. Pada teknik ini digunakan tingkat kesukaran yang dinyatakan dalam skala delta. Apabila pola jawaban individu sesuai dengan pola jawaban pada kelompok maka sekor individu akan dinyatakan wajar. Sebaliknya, apabila pola sekor yang diperoleh individu tidak sesuai dengan pola sekor kelompok maka sekor individu dinyatakan tidak wajar. Nilai negatif pada indeks menunjukkan bahwa, pada peserta ujian tersebut tidak terlihat adanya perbedaan kemampuan untuk menyelesaikan butir soal yang mudah dan sukar. Sebaliknya jika indeks person biserial semakin besar, maka pada peserta tersebut

terlihat jelas perbedaan kemampuan untuk menyelesaikan butir soal yang mudah dan sukar. Apabila pada peserta ujian tidak terlihat jelas perbedaannya antara butir mudah dan sukar, maka dapat dinyatakan bahwa terdapat ketidakwajaran sekor pada peserta tersebut.

Hasil perhitungan, individu yang terdeteksi mempunyai sekor tidak wajar pada SMA hanya 2,49 persen dari 1363 responden. Dari 15 sekolah yang menjadi sampel hanya 4 sekolah yang terdeteksi mempunyai ketidakwajaran sekor. Dari 34 responden yang terdeteksi mempunyai ketidakwajaran sekor, 16 responden tersebut berasal dari satu sekolah sampel.

Berdasarkan perhitungan dengan teknik *caution index*, hanya 254 peserta (18,64 %) yang dinyatakan mempunyai sekor yang wajar. Untuk indeks rendah dan dapat ditolerir terdeteksi 662 peserta (48,57%) ujian. Selanjutnya 267 peserta (19,59 %) terkategori mempunyai indeks ketidakwajaran sedang. Pada indeks kategori tinggi diperoleh angka 132 peserta (9,68 %). Hanya 41 peserta (3,01 %) yang dinyatakan mempunyai indeks yang sangat tidak wajar dan 7 peserta (0,51 %) mempunyai indeks negatif.

Analisis selanjutnya dengan menggunakan teknik *person-fit statistic* yang didasarkan pada *output Quest*. Analisis hasil UAS mata pelajaran Fisika SMA, diperoleh angka 2,78 persen responden (38 orang) mempunyai sekor tidak wajar. Berdasarkan hasil tersebut, maka dapat dikatakan bahwa jumlah peserta yang mempunyai sekor tidak wajar tergolong rendah.

Berdasarkan hasil analisis butir soal secara empiris menggunakan program Quest dari jawaban UAS mata pelajaran Fisika tahun 2015, terlihat bahwa pada dasarnya telah diperoleh butir-butir tes yang cocok dengan model Rasch. Berdasarkan hal itu, akuntabilitas pengukuran telah dapat dipenuhi sebagaimana yang diharapkan. Hasil penelitian ini juga sangat memudahkan guru ketika ditugasi untuk membantu satuan pendidikan melakukan analisis butir soal untuk memenuhi bukti empiris di lapangan, sehingga dapat meningkatkan kualitas butir soal sampai menjadi soal yang terstandar. Analisis butir soal dalam rangka mengembangkan soal UAS terstandar sangat dibutuhkan mengingat sudah banyak satuan pendidikan SMA yang membuat sendiri soal ujian sekolah. Teknik analisis inipun penting untuk digunakan dalam standarisasi soal tes yang dikembangkan oleh Dinas Pendidikan maupun MGMP sehingga dapat diketahui kualitas soal tersebut setelah dikerjakan oleh siswanya. Informasi dari kualitas butir soal tersebut menjadi masukkan dalam pengembangan perangkat butir soal UAS selanjutnya.

Berdasarkan temuan hasil analisis butir menggunakan Quest, maka di lapangan yang berkaitan dengan penyusunan soal UAS yang terstandar, perlu direalisasikan upaya standarisasi tes berbasis sekolah atau MGMP. Satuan pendidikan harus mensikapi hal ini sebagai suatu bentuk otonomisasi dalam penilaian

peserta didik sebagaimana dituntut di dalam Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 3 Tahun 2013 tentang kriteria kelulusan peserta didik dari satuan pendidikan dan penyelenggaraan ujian sekolah. Terlebih dengan adanya kebijakan penyelenggaraan penerimaan mahasiswa baru oleh perguruan tinggi negeri yang menggunakan Pangkalan data Sekolah yang berisi nilai perolehan siswa berdasarkan data hasil ujian oleh sekolah yang bersangkutan.

Dalam penggunaan program analisis butir seperti Quest, kesulitan memahami menu beserta pemanfaatannya ketika melakukan konversi merupakan permasalahan yang biasa terjadi di kalangan pendidik. Kondisi dapat diatasi dengan cara mengintensifkan pemanfaatan program analisis data hasil ujian untuk memperoleh bukti secara empiris, kalau perlu dicanangkan sebagai suatu tuntutan yang harus dipenuhi oleh setiap satuan pendidikan. Kebijakan yang mulai dilaksanakan pada tahun 2015, bahwa setiap satuan pendidikan wajib melaksanakan ujian sekolah untuk mata pelajaran yang diujikan secara nasional akan menjadi momen penting untuk dijadikan modal bagi satuan pendidikan untuk memberikan bukti empiris atas kualitas tes yang diujikan melalui ujian sekolah.

Diskusi selanjutnya adalah berkaitan dengan indeks ketidakwajaran soal, yang juga penting untuk bahan pembuatan kebijakan tentang pelaksanaan ujian sekolah. Dari hasil analisis butir soal diperoleh bahwa *caution index* dari Sato terdeteksi 11,56 persen sample mempunyai ketidakwajaran, dan 18,72 persen sample yang diambil mempunyai tingkat ketidakwajaran yang tinggi. Hal ini secara sederhana menunjukkan bahwa dengan teknik korelasi person biserial, skor peserta dinyatakan tidak wajar apabila memperoleh indeks negatif. Namun jika dibandingkan dengan korelasi point biserial atau korelasi point biserial untuk butir soal, maka selayaknya peserta yang mendapat indeks mendekati nol juga harus dinyatakan merupakan skor tidak wajar. Sebagai pembanding dapat kita lihat indeks daya beda butir soal yang dinyatakan oleh Ebel (dalam Crocker dan Algina, 1986). Untuk klasifikasi yang digunakan peneliti tentang *caution index*, diperlukan analisis yang lebih mendalam dengan cara melakukan *replikasi*, sehingga dapat dilihat sejauh mana klasifikasi tersebut efektif. Naga (1992) menyatakan tentang pendapat para ahli yang menyatakan bahwa apabila indeks semakin besar maka skor tersebut semakin tidak wajar dan sebaliknya apabila semakin kecil maka skor tersebut semakin wajar.

Berdasarkan hasil analisis menggunakan *personal biserial* dan *caution index*, tentang ketidakwajaran skor terlihat bahwa antara grup satuan pendidikan tidak terjadi perbedaan yang signifikan. Namun pada ada beberapa sekolah, yang indeks ketidakwajaran sangat berbeda dengan sekolah yang lain. Ada sekolah yang respondennya mempunyai indeks ketidakwajaran yang cukup tinggi.

Sebaliknya ada sekolah yang mempunyai peserta ujian relatif sedikit terindikasi ketidakwajaran sekornya. Hal ini memperlihatkan bahwa kebijakan dan *fairness* setiap sekolah berbeda tentang pelaksanaan ujian sekolah.

BAB VI

RENCANA TAHAPAN BERIKUTNYA

Sesuai dengan rancangan penelitian semula bahwa pada tahap pertama tahun kedua penelitian telah dihasilkan produk berupa CMI (*Computer Management Instructional*) yang telah dideseminasikan dalam pengembangan assessment terstandar. Tahap pertama sampai pada tahapan ini program awal sudah berhasil direalisasikan dan masuk pada tahap perbaikan dan pengujian terbatas. Tahapan yang sedang dilakukan adalah mengembangkan bank soal tes terstandar agar dapat dimanfaatkan oleh sekolah sekolah di Indonesia lhususnya pada penelitian ini fokus pada soal Fisika SMA.

Tahapan pekerjaan berikutnya dari penelitian ini adalah; (1) melakukan pengujian validitas dan reliabilitas item tes anchor baik secara teoritik (content validity, construct validity dan face validity) dengan melibatkan reviewer ahli, (2) pengujian empirik butir tes di Bank soal dengan menggunakan analisis butir soal (item respons theory), menggunakan program Quest, (3) proses pengurusan hak cipta program CMI, (4) pelatihan guru-guru lebih luas di sekolah mitra dengan melibatkan guru-guru. Hal ini dilakukan sebagai bagian dari keseluruhan tujuan khususnya yaitu; (1) mengembangkan tes terstandar sesuai dengan kompetensi inti dari kurikulum 2013 yang dapat digunakan sebagai perangkat equiting (penyetaraan) dalam penyusunan tes di seluruh sekolah sehingga dapat digunakan sebagai penentu kelulusan dalam sistem ujian akhir nasional yang berkeadilan, (2) mengembangkan model assessment terstandar berbasis CMI (*Computer Management Instructional*) untuk menjamin kesetaraan kualitas penilaian di seluruh sekolah sebagai basis data penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan, (3) menemukan ukuran sampel minimum, pengaruh panjang tes, panjang tes *anchor* minimum, dan metode enyetaraan tes dalam penyetaraan vertikal model kredit parsial menggunakan *common-item nonequivalent groups design*, (4) menghasilkan basis data untuk digunakan dalam penentuan kelulusan dalam sistem ujian akhir nasional dan SNMPTN jalur undangan yang berkeadilan.

BAB VII

KESIMPULAN DAN SARAN

A. Kesimpulan

Pada kebanyakan program tes dalam skala besar, penyusunan tes-tes yang setara merupakan kegiatan yang sangat penting. Hal ini harus dilakukan untuk penanganan yang cepat apabila terjadi kebocoran tes dan untuk membandingkan hasil tes peserta yang menggunakan tes-tes yang berbeda tersebut. Kegiatan ini dapat dilakukan dengan menggunakan teori responsi butir (*item response theory*). Karena digunakan dalam skala luas maka pemanfaatan teknologi *computer mangement system* (CMI) telah memberikan peluang bagi sekolah, guru dan siswa untuk dapat berinteraksi dengan mengakses fasilitas server, desktop maupun aplikasi secara virtual tanpa harus melakukan investasi dan pemeliharaan secara mandiri. Hal ini menjadi suatu peluang yang makin mudah untuk dilakukan seiring perkembangan jaringan data yang semakin variatif dan tersebar luas.

Luaran penelitian yang telah dihasilkan adalah; (1) Software CMI Sistem Informasi Penilaian SMA (SIPSMA), (2) Manual user book untuk program CMI SIPSMA, (3) *Proceeding of International Conference On Research, Implementation And Education Of Mathematics And Sciences (ICRIEMS) 2015, Yogyakarta State University, 17-19 May 2015* (ISBN : 978-979-96880-8-8, lampiran 1), (4) *Proceeding of International Conference On Educational Reseach an Innovation (ICERI), Institute of Research and Community Services Yogyakarta State University, 6-7 May 2015* (ISSN: 2443-1753, dan publikasi jurnal nasional terakreditasi/ jurnal internasional; (1) *Aceptance Jurnal Kependidikan* (LPPM UNY), Jurnal Nasional Terakreditasi (2) proses review Jurnal Ilmu Pendidikan (Universitas Negeri Malang), Jurnal Nasional terakreditasi dan (3) Submit jurnal internasional di *International Journal of Education and Learning* (Canadian Centre of Science and Education).

Hasil analisis menunjukkan bahwa kualitas soal di masing masing wilayah memiliki tingkat kesukaran yang berbeda. Paket soal di wilayah Jawa Barat paling tinggi dibandingkan paket soal di wilayah lain. Persamaan linear equating yang terbaik terhadap skor akhir yang tidak merugikan peserta ujian adalah penyetaraan dari paket sukar ke paket mudah dengan persamaan; (1) $Y' = 1,543 X - 3.976$, untuk equating dari Paket A ke B (2) $Y' = 1,876 X - 14,342$, untuk equating dari Paket

B ke C (3) $Y' = 1,443 X - 3,203$, untuk equating dari Paket C ke A , (4) $Y' = Y' = 1,384 X - 0.932$ untuk equating dari Paket D ke A , (5) $Y' = 1,214 X - 0,607$ untuk equating dari Paket A ke C, (6) $Y' = 1,229 X - 4,987$ untuk equating dari Paket B ke C, dan (7) $Y' = 1,256 X - 1,660$. untuk equating dari Paket B ke D.

Penelitian ini juga mendeteksi ketidakwajaran sekor dengan korelasi person biserial (*person-fit statistic*) dan indeks kehati-hatian (*caution index*) dari Sato. Paket program yang digunakan untuk melakukan analisis butir adalah Quest, dengan elemen sentral *Rasch Model* (RM). Dari hasil analisis butir soal diperoleh bahwa *coution index* dari Sato terdeteksi 11,56 persen sample mempunyai ketidakwajaran, dan 18,72 persen sample yang diambil mempunyai tingkat ketidakwajaran yang tinggi.

B. Ucapan Terimakasih

Model Assessment Terstandar Berbasis CMI (*Computer Management Instructional*) baru dapat dikembangkan untuk tingkat SMA, sehingga pengembangannya masih memerlukan review dan uji coba yang lebih baik, mengingat masih sangat beragamnya kualitas sekolah di wilayah negara kesatuan Republik Indonesia. Kami berharap sekecil apapun kontribusi yang dapat diberikan tetap dapat memberikan manfaat bagi penelitian-penelitian berikutnya. Tidak lupa terima kasih kepada Direktorat Penelitian dan Pengabdian Pada Masyarakat (Ditlitabmas) Dirjen DIKTI yang telah mendanai penelitian ini melalui skim Hibah Kompetensi sehingga penelitian ini terlaksana

DAFTAR PUSTAKA

- Allen, Mary J., dan Wendy M Yen. 1989. *Introduction to Measurement Theory*. California: Cole Publishing Company.
- Angoff, W. H. 1982. Uses of Difficulty and Discrimination Indices for Detecting Item Bias In RA Berk. *Handbook of Methods for Detecting Item Bias*. Baltimore: John Hopkins University Press.
- Asmin. 2004. Implementasi Teori Responsi Butir dan Fungsi Informasi Butir Tes dalam Pengujian Hasil Belajar Akhir di Sekolah. *Jurnal Pendidikan dan Kebudayaan*, X (48): 234-245.
- Barnard. John. J. 1996. *In Search for Equity in Educational Measurement: Traditional Versus Modern Equating Methods*. Makalah: Disampaikan pada ASEESA National Conference di HSRC Conference Centre. Pretoria: Afrika Selatan.
- Camilli, Gregory, dan Lorrie A. Shepard. 1994. *Methods for Identifying Biased Test Items*. California: Sage Publication.
- Chong Ho Yu dan Sharon E. Osborn. 2005. Test Equating by Common Items and Common Subject: Concepts and Applications. *Practical Assessment, Research & Evaluation*. X (4): 187-198.
- Crocker, Linda, & Algina, James. 1986. *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston, Inc.
- Djaali. 2004. *Pengukuran Dalam Bidang Pendidikan*. Jakarta: Program Pascasarjana Universitas Negeri Jakarta.
- Dorans, N. J. (2004). Equating, concordance, and expectation. *Applied Psychological Measurement*, 28 (4), 227-246.
- Gronlund, Norman. E. 1985. *Measurement and Evaluation in Teaching*. New York: Macmillan Publishing Company.
- Hambleton, Ronald K, Swaminathan, H., dan Jane Rogers, H. 1991. *Fundamentals of Item Response Theory*. London: Sage Publications.
- Hambleton, Ronald K., dan Swaminathan, H. 1985. *Item Response Theory: Principle and Applications*. Boston: Kluwer Nijhoff Publishing.
- Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Journal of Educational measurement* (4th ed., pp. 187{220). Westport, CT: Greenwood.
- Jihad, Asep, Abdul Haris. 2011. *Evaluasi Pembelajaran*. Multi Pressindo: Yogyakarta.

- Kim, S., von Davier, A. A., & Haberman, S. (2008). Small-sample equating using a synthetic linking function. *Journal of Educational Measurement*, 45, 325{342}
- Kolen, Michael J., dan Robert L. Brennan. 2004. *Test Equating, Scaling, and Linking: Methods and Practices*. New York: Springer.
- Kolen, Michael J., dan Robert L. Brennan. 1995. *Test Equating*. New York: Springer Verlag.
- Kumaidi. 2000. Standardisasi Butir Soal. *Jurnal Pendidikan dan Kebudayaan*. V(5): 132-143.
- Livingstone, S. A., Doran, N. J. dan Wright, N. K. 1990. What Combination of Sampling and Equating Methods Work Best?. *Applied Measurement in Education*. III (2): 73-95.
- Livingston, S. A., & Kim, S. (2009). The circle-arc method for equating in small samples. *Journal of Educational Measurement*, 46, 330{343}
- Lord, F. M. (2009). The standard error of equipercentile equating. *Journal of Educational Statistics*, 7, 165{174}
- Lord, Frederick, M. 1990. *Applications of Item Response Theory to Practical Testing Problems*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Mary J. Allen and Wendy M Yen, 1989, *Introduction to Measurement Theory*, California: Broke.
- McDonald, Roderick P. 1991. *Test Theory: A Unified Treatment*. New Jersey: Lawrence Erlbaum Associates Publisher.
- Naga, Dali, S. 1992. *Pengantar Teori Sekor Pada Pengukuran Pendidikan*. Jakarta: Besbats.
- Nitko, Anthony. J. 1992. *Criterion Reference Testing Workshop: Handouts and Reading Material Tidak dipublikasikan*. Cipayung, Bogor: Examination Development Unit (Puslitbang Sisjian).
- Miyatun, Erna., dan Djemari Mardapi. 2000. Komparasi Metode Penyetaraan Tes Menurut Teori Responsi Butir. *Jurnal Penelitian dan Evaluasi*. II (3): 124-132.
- Peraturan Pemerintah No. 19 Th 2005 Tentang: Standar Nasional Pendidikan (SNP). Bandung: Citra Umbara.
- Peterson, N.S., Kolen, M.J., dan Hoover, H.D. 1989. Scaling, Norming, and Equating. In R.L. Linn (Ed), *Educational Measurement*. New York: Macmillan.
- Rahayu, Wardani. 2008. Pengaruh Metode Linking Terhadap Banyak Butir False Positive pada Pendeteksian DIF Berdasarkan Teori Responsi Butir. *Disertasi*. Jakarta: Universitas Negeri Jakarta.

- Ridho, Ali. 2007. Karakteristik Psikometrik Tes Berdasarkan Pendekatan Teori Tes Klasik dan Teori Respon Aitem. *Jurnal Insan Media*. II (2): 1-28.
- Setiadi, Hari. 1998. Bank Soal yang Dikalibrasi dengan Konsep IRT Memecahkan Permasalahan Ujian-ujian Sistematis yang Diadakan pada Periode-periode Tertentu, *Jurnal Kajian Dikbud* IV (13).
- Setiadi, Hari. 2009. Permasalahan dan Solusinya dalam Pelaksanaan Ujian Nasional di Masa Mendatang, *Matahari: Jurnal Penelitian dan Pendidikan*. X (1): 66-74.
- Skaggs, G. (2005). Accuracy of random groups equating with very small samples. *Journal of Educational Measurement*, 42, 309-330.
- Susongko, Purwo. 2005. Penyetaraan Parameter Butir Secara Konkuren untuk Menguji Secara Statistik Keberadaan Item Function (DIF). Makalah: Disampaikan pada Seminar Nasional *Hasil Penelitian tentang Evaluasi Hasil Belajar serta Pengelolaannya*. Pascasarjana UNY Didukung oleh Direktorat P2TK & KPT dan HEPI, Yogyakarta, 14-15 Mei 2005.
- Sukirno, D. S. 2007. Penyetaraan Tes UAN: Mengapa dan Bagaimana. *Jurnal Cakrawala Pendidikan*. XXVI (3): 305-321.
- Syarifah. 2007. Persyaratan Analisis Instrumen Sebagai Prasyarat Ketepatan Hasil Analisis Dalam Penelitian Pendidikan. *Cakrawala Pendidikan*. XXVI (2): 15-27.
- Swediati, Nonny. 1997. *Metode untuk Penyetaraan (Equating) Sekor Tes Secara Klasik*. Pusat Pengujian Balitbang Dikbud: Jakarta.
- Wibowo, Mungin Eddy. 2011. Kondisi Psikologis Siswa dalam Menghadapi Ujian Nasional, *Buletin BNSP: Media Komunikasi dan Dialog Standar Pendidikan*. VI (1): 7-11.
- Widhiarso, Wahyu. 2011. Aplikasi Teori Respon Butir untuk Menguji Invariansi Pengukuran Psikologis Guna Keperluan Survei dan Seleksi Pekerjaan. *Jurnal Psikobuana*. III (2): 104-117.
- Tumilisar, A.V.J. 2006. Akurasi Relatif Penyetaraan Sekor Tes untuk Sampel Berukuran 300 Ditinjau dari Metode Penyetaraan dan Teknik Penghalusan. *Jurnal Pendidikan Penabur*. V (6): 1-19.
- Zhu, W. 1998. Test Equating: What, Why, How?. *Research Quarterly for Exercises and Sport*. Wayne State University.

